(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
8 February 2007 (08.02.2007)

PCT

(10) International Publication Number
WO 2007/016344 A1

(51) International Patent Classification:
G01N 35/10 (2006.01) *G01N 21/53* (2006.01)
G01N 35/08 (2006.01)

(21) International Application Number:
PCT/US2006/029413

(22) International Filing Date: 27 July 2006 (27.07.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/703,814 28 July 2005 (28.07.2005) US

(71) Applicant *(for all designated States except US)*: **THE GOVERNMENT OF THE UNITED STATES OF AMERICA as represented by the Secretary, Department of health and Human services** [US/US]; 6011 EXECUTIVE BOULEVARD, Suite 325, Rockville, MD 20852 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **MINTON, Allen, P.** [US/US]; 9205 Bardon Road, Bethesda, MD 20814 (US). **ATTRI, Arun, K.** [IN/IN]; 1/30 Sunder Vihar, New Dehli 110 087 (IN).

(74) Agent: **KOWALCHYK, Katherine, M.**; MERCHANT & GOULD P.C., P.o.box 2903, Minneapolis, MN 55402-0903 (US).

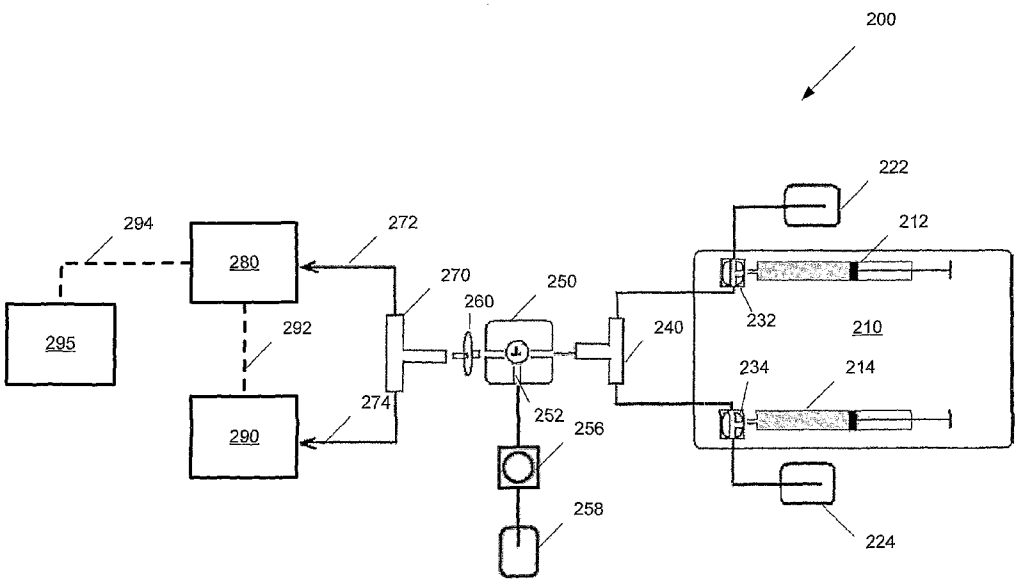
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

[Continued on next page]

(54) **Title:** DETECTING AND CHARACTERIZING MACROMOLECULAR INTERACTIONS IN A SOLUTION WITH A SIMULTANEOUS MEASUREMENT OF LIGHT SCATTERING AND CONCENTRATION



(57) **Abstract:** Systems and methods for detecting macromolecular interactions in solution. Systems can include a dispenser module to dispense a solution including a macromolecule, a detector to measure a light scattering associated with the macromolecule in the solution, and to measure a concentration associated with the macromolecule in the solution. In some embodiments, a first detector and a second detector can be positioned in parallel, so that the first and second detectors take simultaneous measurements of light scattering and light absorbance. Example methods can be used to analyze the data to detect and model self- and hetero-associations of the macromolecule.

WO 2007/016344 A1

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

DETECTING AND CHARACTERIZING MACROMOLECULAR INTERACTIONS IN A SOLUTION WITH A SIMULTANEOUS MEASUREMENT OF LIGHT SCATTERING AND CONCENTRATION

This application is being filed on 27 July 2006, as a PCT International Patent application in the name of The Government of the United States of America as represented by the Secretary, Department of Health and Human Services, applicant for the designation of all countries except the US, and Allen P. Minton, a citizen of the U.S., and Arun Attri, a citizen of India, applicants for the designation of the US only, and claims priority to U.S. Provisional Patent Application No. 60/703,814, filed July 28, 2005.

STATEMENT REGARDING FEDERALLY SPONSORED RESEARCH OR DEVELOPMENT

This invention has been developed with the support of the Department of Health and Human Services. The Government of the United States of America has certain rights in the invention disclosed and claimed herein below.

COPYRIGHT NOTICE

TECHNICAL FIELD

This invention relates to systems and methods for detecting macromolecular interactions in solution.

BACKGROUND

The detection of protein-protein interactions in solution is important for analyzing the structure and function of proteins. Such analysis assists in the understanding of how complex biochemical systems function in response to changes in composition and environment. There are a variety of systems and methods for studying such protein interactions. For example, high-throughput assays, such as the yeast two-hybrid and tandem pull-down assays, provide qualitative information

about strong interactions. However, such methods provide little information regarding weaker interactions and reversible associations involved in a regulatory process.

Other methods such as physical-chemical techniques provide high-resolution information about association equilibria. Examples of these types of methods include sedimentation equilibrium, isothermal titration caloriety, osmotic pressure, and a variety of spectroscopic assays. For example, the sedimentation equilibrium technique provides information about the composition dependence of a signal-average buoyant mass. The observed dependency can then be modeled in the context of schemes for association.

Another high-resolution technique involves the use of static light scattering for determining molar masses and radii of gyration of macromolecules. For example, the composition dependence of the light scattering of a mixture of macrosolutes can be analyzed using a batch procedure to yield information about interactions between macrosolute species. A series of solutions containing a macrosolute at different concentrations can be prepared, and each of the series of solutions can be analyzed in sequence in a scattering cell at multiple angles. The relative apparent weight-average molar masses and/or z-average radius of gyration can then be calculated by linear regression for each solution, and solute-solute interaction can be identified as a concentration dependence of the molar mass.

Although processes such as sedimentation and light scattering can provide high resolution data, there are distinct disadvantageous because such the processes are low through-put and time/labor intensive.

It is therefore desirable to provide systems and methods to measure macromolecular interactions that generate data of increased precision, sensitivity, range, and/or application. It is also desirable to provide systems and methods for maximizing through-put and/or minimizing manual intervention.

SUMMARY

This invention relates to systems and methods for detecting and characterizing macromolecular interactions in a homogenous or heterogeneous solution of macromolecules. Examples of macromolecules include proteins, DNA, RNA, biopolymers, organic polymers, and inorganic polymers.

In some embodiments, systems can include a dispenser module to dispense at least one solution including at least one macromolecule, a detector to measure a light scattering associated with the macromolecule in the solution, and to measure the concentration of macromolecule in the solution. In some embodiments, a first and second detector can be positioned in parallel, so that the first and second detectors take simultaneous measurements of light scattering and concentration. In some embodiments, methods can be used to analyze the data to detect and model self- and hetero-associations of the macromolecule. These methods can identify complexes and evaluate equilibrium constants for hetero-associations.

BRIEF DESCRIPTION OF THE DRAWINGS

Reference will now be made to the accompanying drawings, which are not necessarily drawn to scale, and wherein:

Figure 1 is a schematic representation of one embodiment of a system for detecting macromolecular interactions;

Figure 2 is a schematic representation of another embodiment of a system for detecting macromolecular interactions;

Figure 3 is an example graph showing example data sets of light scattering and absorbance of ovalbumin plotted as a function of time, the two curves representing each data set overlap one another;

Figure 4A is an example graph showing further refinement of the data shown in Figure 3;

Figure 4B is an example graph showing the best-fit residuals of the data shown in Figure 4A;

Figure 5A is an example graph showing further refinement of a filtered subset of the data shown in Figure 3;

Figure 5B is an example graph showing the best-fit residuals of the data shown in Figure 5A;

Figure 6 is an example graph showing calculated molar masses for example proteins including: (1) fibrinogen; (2) alcohol dehydrogenase; (3) bovine serum albumin (nonequilibrium mixture of monomer + oligomers); (4) hemoglobin; (5) bovine serum albumin; (6) ovalbumin; (7) pepsinogen; (8) β-lactoglobulin (mixture of A and B); (9) β-lactoglobulin A; (10) chymotrypsinogen A; (11) lysozyme; and (12) cytochrome *c*.;

Figure 7 is an example graph showing experimentally measured values of $\langle R \rangle/K$ for β -lactoglobulin;

Figures 8A, 8B, and 8C are examples graphs showing alternative 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 $w_{B,tot}$.

Figure 9A is an example graph showing an example data set of light scattering plotted as a function of time for solutions of chymotrypsin and trypsin inhibitor;

Figure 9B is an example graph showing an example data set of light absorbance plotted as a function of time for solutions of chymotrypsin and trypsin inhibitor;

Figure 10 is an example graph showing a scaled Rayleigh ratio $R(\theta,t)/K$ calculated from the gradient shown in Figures 9A and 9B plotted against $\sin^2(\theta/2)$ and $A_{280}(t)$; and

Figure 11 is an example graph showing a value of $\langle R \rangle/K$ calculated from the data plotted in Figures 9A and 9B plotted as a function of f_A, along with calculated best fits of three models. Lower panel shows best fit of simple 1-1 association model (dotted curve) and the best fit of the equilibrium model (dashed curve). The solid curve in the lower panel represents the best fit of the relaxed equilibrium model. Upper panel shows best fit residuals of the constrained (dashed curve) and relaxed (solid curve) models.

Figure 12 is a schematic representation of another embodiment of a system for detecting macromolecular interactions.

Figures 13A-13C are example graphs showing $\langle R \rangle/K$ plotted as a function of f_A for a composition gradient of chymotrypsin and bovine pancreatic trypsin inhibitor at pH 4.4 (Fig. 13A), pH 5.4 (Fig. 13B), and pH 8.0 (Fig. 13C).

Figures 14A and 14B are example graphs showing the contribution of individual species to the composition gradient scattering profile of chymotrypsin and BPTI using parameters obtained from an equilibrium model described by Eqs. 20-24 to data obtained at pH 5.4. Fig. 14A shows the calculated concentrations of individual species. Each line represents an individual species: chymotrypsin (A), BPTI (B), hetero-association of chymotrypsin and BPTI (AB), self-association of chymotrypsin (A_2). Fig. 14B shows the calculated concentrations of individual species (A, B, and AB) to the total scattering profile (total).

Figures 15A and 15B are example graphs showing the best-fit values of equilibrium association constants characterizing hetero-association of chymotrypsin and BPTI (Fig. 15A) and self-association of chymotrypsin (Fig. 15B) plotted as a function of pH. In Fig. 15A, circles represent values obtained by modeling a single composition gradient experiment, squares represent values obtained by global modeling of a composition gradient together with one or two dilution experiments conducted on individual proteins, diamonds represent values reported by Vincent and Lazdunski (1973, *Eur. J. Biochem.*, 38:365-372) calculated from the ratio of measured association and measured rate constants, and triangles represent values reported by Rigbi as quoted in Vincent and Lazdunski. In Fig. 15B, triangles represent values obtained by modeling a single composition gradient experiment, squares represent values obtained by modeling a single dilution experiment conducted on pure chymotrypsin, circles represent values obtained by global modeling of composition gradient and dilution experiments, and diamonds represent values reported by Aune and Timasheff (1971, *Biochem.*, 10:1609-1617).

Figure 16 is an example graph showing concentration-dependent scattering of FtsZ as a function of total protein concentration. Open circles represent experimental data. The curve in was calculated from an inverse-decay model using any of several sets of correlated parameter values leading to identical fits of the data.

Figures 17A and 17B are examples graphs showing K_i (Fig. 17A) and G_i (Fig. 17B) for addition of monomer to form *i*-mer, plotted as a value of *i*. Open circles in Figs. 17A and 17B were calculated using the experimentally observed dependence of $<R>/K$ on w_{tot}. The curve in Fig. 17A and 17B was calculated using the parameters obtained by modeling sedimentation equilibrium data as described in Rivas et al., 2000. *J. Biol. Chem.*, 275:11740-11749.

Figure 18A is an example graph showing an example data set of light scattering plotted as a function of time for ascending and descending gradients of concentration of bovine serum albumin.

Figure 18B is an example graph showing an example data set of differential refractive index plotted as a function of time for ascending and descending gradients of concentration of bovine serum albumin.

DETAILED DESCRIPTION

I. Definitions

As used herein, "macromolecule" refers to a molecule of high relative molecular mass. Non-limiting examples of a macromolecule include a biopolymer, organic polymer, inorganic polymer, or copolymer thereof.

The term "biopolymer" includes polypeptides such as proteins, receptors, antibodies, antibody fragments, monobodies, and immunoadhesions, polynucleotides such as DNA and RNA, starches, lipids, cellulose, lignans, and the like. Examples of proteins include, but are not limited to fibrinogen, alcohol dehydrogenase, bovine serum albumin (nonequilibrium mixture of monomer + oligomers), hemoglobin, bovine serum albumin, ovalbumin, pepsinogen, β-lactoglobulin (mixture of A and B), β-lactoglobulin A, chymotrypsin, chymotrypsinogen A, lysozyme, and cytochrome *c*.

The term "organic polymer" includes polyamide, polyethylene, polylactate, polyacrylate, polyolefin, polyglycolate, polypropylene, polystyrene, polyvinylchloride, fluoropolymers, polymethylmethacrylate, polyethyleneterephthalate, copolymers thereof, and the like.

The term "inorganic polymer" includes polysiloxanes, polysilanes, polygermanes, polystannanes, polyphosphazenes, copolymers thereof, and the like.

The term "balanced" or "balanced flow" as used herein refers to a composition of the solution flowing in the concentration detector at a particular point in time corresponding to the composition of the solution flowing into the light scattering detector at the same or substantially the same point in time.

The term " simultaneous" as used herein refers to the collection of concentration and scattering data from the identical element of volume or two elements of volume with the same solute composition, as the composition of the sample is gradually being varied with time. In some embodiments, the concentration and light scattering data are collected at the same or substantially the same point in time from one or more detectors.

II. Methods and Systems of Carrying Out the Invention

Quantitative characterization of reversible macromolecular associations between different species of biological macromolecules in solution assists in the understanding of how complex biochemical systems respond to changes in

composition and environmental variables. The composition dependence of the light scattering of a mixture of macromolecules can be analyzed to yield information about attractive and repulsive interactions between individual macromolecule species. However, acquisition of such information utilizing conventional batch procedures is a time-consuming and labor-intensive process; hence it is rarely utilized.

Embodiments of the present invention relate to systems and methods for detecting and characterizing reversible macromolecular interactions in a homogenous or heterogeneous solution of macromolecules. Utilizing a novel analytical procedure, the data acquired by the system can be interpreted rapidly to yield reliable estimates of the molar mass(es) of macromolecule species and the strength of reversible associations between them.

The systems of the invention can detect and characterize self-associations of macromolecules in a homogenous solution. For example, the systems of the invention can detect and characterize macromolecular interactions such as monomer, dimer, or trimer formation, and the like of macromolecules, such as proteins, in a homogenous solution.

The systems of the invention can also detect macromolecular interactions between one or more macromolecule species in a heterogeneous solution. For example, the systems of the invention can detect and characterize macromolecule interactions between a protein and DNA or RNA, a protein and an antibody, two or more different species of proteins, a protein and an organic polymer, inorganic polymer, or biopolymer, two or more species of organic polymers, two or more species of biopolymers, a biopolymer and an organic polymer, a biopolymer and an inorganic polymer, or a organic polymer and an inorganic polymer. In an embodiment, the systems of the invention can detect macromolecular interactions between a protein and an agonist or antagonist. The agonist or antagonist can be a protein, antibody, antibody fragment monobody, immunoadhesion, or receptor.

The systems and methods of the invention are useful, inter alia, in methods for determining the extent of aggregation of a particular macromolecule such as a protein. In some embodiments of pharmaceutical formulations, aggregation of the biologically active agent can greatly decrease efficacy and/or increase toxicity. Alternatively, the methods and systems of the invention can be used, inter alia, to measure and compare the strengths of binding interaction between a macromolecule

and a number of different binding partners. Such methods would allow identification of binding partners having a desired level of binding affinity.

The systems of the invention generally include a dispenser module, a light scattering detector, and a concentration detector. The dispenser module generally includes a mixer and one or more solute reservoirs for solutions of macromolecules. In an embodiment, the macromolecule solutions are preferably at least 95% pure using conventional purification methods. The systems of the invention can theoretically analyze any number of macromolecular species in a solution, however, preferably, about 1 to 4 different macromolecular species can be analyzed. In an embodiment, the dispenser module comprises at least two solute reservoirs. In another embodiment, the dispenser module comprises at least three solute reservoirs. In yet another embodiment, the dispenser module comprises at least four solute reservoirs. The dispenser module can optionally include a reservoir for solvent.

The dispenser module is configured to dispense a solution stream comprising a time-varying composition of one or more macromolecule species. In an embodiment, the dispenser module provides a stepwise upward or downward gradient of solute concentration that varies roughly linearly over a period of time. In an embodiment, as little as 1 ml of solvent and 1 ml of a stock macromolecule solution having an absorbance greater than 0.1 OD units at the selected wavelength is sufficient to provide the gradient. In an embodiment, the dispenser module comprises a robotic element that sequentially introduces multiple samples.

In an embodiment, the output of the dispenser module is connected to at least one flow cell comprising a light scattering detector and a concentration detector. In some embodiments, a single flow cell may be utilized. In other embodiments, more than one flow cell may be utilized. When more than one flow cell is utilized, the output of the dispenser module is connected to a splitter. The splitter splits the solution stream into parallel streams with a similar or substantially similar flow rate. One of the parallel streams flows into a light scattering detector. The other parallel stream flows into a concentration detector. In an embodiment, the flow rate of the stream flowing into a detector is dependent on the size of the detector's flow cell. In an embodiment, the flow rate is within the flow rate parameters of the detector. As smaller flow cells and detectors become available, smaller sample volumes and/or flow rates can be used to generate the data. In an embodiment, the flow rate can be from about 0.1 ml/min to about 2 ml/min. In another embodiment, the flow rate can

be from about 0.75 ml/min to about 1.25 ml/min. In yet another embodiment, the flow rate is about 1 ml/min.

In some embodiments, the composition of the solution flowing into the concentration detector and light scattering detector is "balanced", meaning the composition of the solution flowing in the concentration detector at a particular point in time corresponds to the composition of the solution flowing into the light scattering detector at the same point in time. In such an embodiment, balanced flow through the detectors can be achieved by calibrating the flow rate of the stream(s) flowing into the concentration detector and light scattering detector with a solution comprising a macromolecule that does not self-associate. The non-associating macromolecule should be large enough to provide a clean scattering signal. In an embodiment, the macromolecule has a molecular weight of at least 20,000 daltons. Examples of non-associating proteins or synthetic macromolecules that are non-associating include ovalbumin, serum albumin, and starburst dendromers. The flow rate of the stream(s) can be calibrated by running the non-self associating protein or synthetic macromolecule solution through the system and collecting signal intensities at a plurality of data points. The signal intensity from each of the collected data points is scaled to relative units and plotted against the data points to form a signal intensity curve for each of the detectors. The flow rate of the stream flowing into each of the detectors is adjusted until the signal intensity from each detector as a function of time is approximately proportional. See, for example, Figure 3, which illustrates an example of scaled light scattering and concentration data of ovalbumin plotted as a function of time. Figure 3 shows 2 curves (one for the concentration detection and one for the light scattering detection) that are superimposed on one another.

The light scattering detector can be selected to measure the scattering of light from a plurality of angles. In an embodiment, the light scattering detector measures the scattering of light from at least 15 different angles. One example of a multiangle light scattering detector is a DAWN-EOS multiangle light scattering detector manufactured by Wyatt Technology Corporation of Santa Barbara California. Other types of multiangle light scattering detectors are known. The concentration detector can be selected to measure, for example, absorbance or refraction of light. Data from the light scattering detector and concentration detector is recorded at regular

intervals and this data is analyzed to detect and model self-associations and hetero-associations of macromolecule species in the solution stream.

Embodiments of the present invention will now be described more fully hereinafter with reference to the accompanying drawings. The representative embodiments described hereafter relate to detecting and characterizing macromolecular interactions of one or more species of protein in a solution using the systems and methods of the invention. Principles associated with this invention may, however, apply to the detection and characterization of macromolecular interactions of any macromolecule species or combination thereof in a solution. Principles associated with this invention can be embodied in many different forms and should not be construed as limited to the embodiments set forth herein. Instead, these embodiments are provided so that this disclosure will be thorough and complete, and will fully convey principles of the invention to those skilled in the art. Like numbers refer to like elements throughout.

A. Instrumentation

Referring now to Figure 1, a system 100 for detecting macromolecular interactions is shown. System 100 generally includes a dispenser 110, a splitter 170, and detectors 180 and 190.

Dispenser 110 is configured to dispense a solution stream including one or more macromolecules. In an embodiment, dispenser 110 is configured to vary over time the concentration of the macromolecule in the solution. In an embodiment, dispenser 110 is configured to dispense a solution stream comprising two or more different species of macromolecules. In another embodiment, dispenser 110 is configured to dispense a solution stream comprising three of more different species of macromolecules. In yet another embodiment, dispenser 110 is configured to dispense a solution stream comprising four or more different species of macromolecules.

The output of dispenser 110 is connected to splitter 170. Splitter 170 splits the solution stream into parallel streams with a similar or substantially identical flow rate. One of the parallel streams is delivered to detector 180. Another of the parallel streams is delivered to detector 190. In the example embodiment, the parallel streams are delivered to each of detectors 180 and 190 at substantially the same time. In an embodiment, the composition of the solution stream flowing into

detector 180 at a particular point in time corresponds to the composition of the solution stream flowing into detector 190 at the same point in time.

Detector 180 is selected to measure an intensity of light scattered by the macromolecule in the solution stream. In an embodiment, detector 180 is configured to measure an intensity of light scattered by the macromolecule in the solution stream at a plurality of angles. In an embodiment, detector 180 is multi-angle light scattering detector. One example of a multi-angle light scattering detector is a DAWN-EOS multi-angle light scattering detector (Wyatt Technology Corporation, Santa Barbara, CA). Other multi-angle light scattering detectors are known.

Detector 190 is selected to measure the concentration of macromolecules in the solution stream. In an embodiment, detector 190 is a light absorbance detector. One example of an absorbance detector is a variable-wavelength UV-visible absorbance detector (Milton Roy SM3100, Thermo Finnegan, West Palm Beach, FL). Other absorbance detectors are known.

In an embodiment, detector 190 is a refractometer. One example of a refractometer is a Leica ARIAS 500 Abbe refractometer (Reichert Instruments, Buffalo, NY). Another example of a refractometer is a differential refractive index detector such as the Bischoff RI8120 (Bischoff Chromatography, Leonburg, Germany) or Waters 2414 refractive index detector (Waters Corporation, Milford, MA). Other refractometers are known.

When measuring concentration changes by refractive index, most proteins have identical refractive increments to within very tight limits of variation. Therefore, in most cases it is not necessary to measure refractive increments independently in order to determine protein concentration. This feature of a differential refractive index detector allows, for example, quantification of a small amount of protein that has not been independently characterized. When analyzing interactions between two macromolecular species that have different refractive increments, such as a protein and a nucleic acid, the refractive increment of any particular hetero-oligomeric species will be the mass average of the refractive coefficient of any hetero-oligomer. In an embodiment, the macromolecular solute is dialyzed against the buffers with which they will be dissolved so that only the macromolecule contributes to the refractive index gradient.

In the illustrated embodiment, since parallel streams of similar flow rate are delivered to each of detectors 180 and 190 at substantially the same time, detectors 180 and 190 each perform measurements on substantially the same concentration of the solution stream for each measurement. Data from detectors 180 and 190 is recorded for analysis as described below.

Referring now to Figure 2, another system 200 for measuring macromolecular interactions is shown. System 200 is similar to system 100 described above, but includes additional components. System 200 generally includes a dispenser 210, solution reservoirs 222, 224, a mixer 240, a valve 250, a filter 260, a splitter 270, detectors 280, 290, and a computer system 295. In an embodiment, dispenser 210 comprises three or more solution reservoirs. In another embodiment, dispenser 210 comprises four or more solution reservoirs.

In the example shown, dispenser 210 includes dual-syringes 212, 214 connected to valves 232, 234, respectively. Solution reservoirs 222, 224 are also connected to three-way valves 232, 234, respectively. In the example shown, reservoir 222 contains stock solution, and reservoir 224 contains a solvent. Valves 232, 234 are programmable so that valves 232, 234 can be switched between a filling mode and a delivery mode. For example, for the filling mode, valve 232 can be switched to allow syringe 212 to be filled with stock solution from reservoir 222, and valve 234 can be switched to allow syringe 214 to be filled with solvent from reservoir 224. Valves 232, 234 can likewise be switched for the delivery mode to dispense fluid from syringes 212, 214.

A rate of delivery from syringes 212, 214 can be controlled. In one example, the rate of delivery from syringes 212, 214 can be controlled to create a gradient in concentration of the solution that is dispensed by dispenser 210 and/or to control the flow rate to provide a balanced flow rate. See, for example, the example source code, provided at the Appendix hereto, written in Turbo Pascal and used to control a Hamilton Microlab 540C dual-syringe precision dispenser manufactured by Hamilton Company of Reno, Nevada.

In example embodiments, dispenser 210 is a Hamilton Microlab 540C or 900 dual-syringe precision dispenser manufactured by Hamilton Company of Reno, Nevada. Other dispenser systems can be used.

Valves 232, 234 are, in turn, connected to T-junction or mixer 240. Mixer 240 combines stock solution from syringe 212 and solvent from syringe 214 into a

single stream of solution ("solution stream"). In one example, mixer 240 is a stream mixer (Upchurch Scientific, Oak Harbor, WA). Other types of mixers can be used.

In the example shown, the solution stream from mixer 240 is delivered to valve 250. Valve 250 can be a three-way value that can be switched between a first mode and a second mode. For example, valve 250 can be switched to the first mode to allow a solution to be introduced into an inlet 252 to, for example, purge system 200. Valve 250 can be switched to the second mode to deliver the solution stream to filter 260.

Alternatively, in some embodiments described below, a reservoir 258 is connected to inlet 252 by a peristaltic pump 256 that is used to pump solution from reservoir 258 into valve 250. In an embodiment, reservoir 258 contains a solvent and reservoirs 222 and 224 each contain stock solution. In an embodiment, the stock solution in reservoir 222 is different from the stock solution in reservoir 224.

Filter 260 is an inline filter that removes particles and other impurities from of the solution stream. In one example, filter 260 is a ANOTOP 0.1-μm filter manufactured by Whatman plc of the United Kingdom. Other types of filters can be used.

Output from filter 260 is delivered to a T-junction or splitter 270. Splitter 270 is similar to that of mixer 240, except splitter 270 is used to split the solution stream for delivery lines 272, 274. In the example shown, delivery lines 272, 274 are parallel lines. The solution stream is split into parallel streams by splitter 270, and the parallel streams are delivered to lines 272, 274. As described further below, the parallel streams in lines 272, 274 are adjusted so as to be similar in flow rate.

Line 272 is connected to deliver one of the parallel streams to detector 280. Detector 280 measures an intensity of light scattered by the macromolecule in the solution stream. In some embodiments, detector 280 is configured to measure light scattering at a plurality of angles. For example, detector 280 can be configured to measure light scattering at up to fifteen angles. In one example, detector 280 is a DAWN-EOS multi-angle light scattering detector including a temperature-regulated K5 flow cell manufactured by Wyatt Technology Corporation of Santa Barbara, California. Other multi-angle light scattering detectors are known.

Line 274 is connected to deliver one of the parallel streams to detector 290. Detector 290 measures the concentration of the macromolecule in the solution stream. In an embodiment, detector 290 is a light absorbance detector. One

example of an absorbance detector is a variable-wavelength UV-visible absorbance detector (Milton Roy SM3100, Thermo Finnegan, West Palm Beach, FL). Other absorbance detectors are known. In an embodiment, detector 290 is a refractometer. One example of a refractometer is a Leica ARIAS 500 Abbe refractometer (Reichert Instruments, Buffalo, NY). Other refractometers are known.

Data from absorbance detector 290 and light scattering detector 280 can be collected substantially simultaneously using ASTRA software, Release No. 4.90.04, from Wyatt Technology Corporation. Because the flow rate of the parallel streams through detectors 280 and 290 is balanced, a measured concentration at a given point in time corresponds to a measured light scattering at the same point in time. System 200 therefore allows for the substantially simultaneous collection of concentration and scattering data from an element of volume or two elements of volume with the same solute composition, as the composition of the sample is gradually being varied with time.

In the example shown, an output of absorbance detector 290 is connected to light scattering detector 280 by connection 292. For example, absorbance detector 290 can output an analog signal (e.g., 1 volt per absorbance unit) to an auxiliary port of light scattering detector 280. Light scattering detector 280 is, in turn, connected to computer system 295 by connection 294. In alternative embodiments, both detectors 280, 290 can be directly connected to computer system 295. In alternative embodiment, for example, light scattering detector 280 can output a signal to an auxiliary port of concentration detector 290. Concentration detector 290 is, in turn, connected to computer system 295. Connections 292, 294 can be wired or wireless connections.

Data recorded from detectors 280, 290 can be communicated to computer system 295 for analysis, as described below. In one example, computer system 295 includes at least one processing unit, memory, and storage. Computer system 295 also contains communications connections that allow the device to communicate with other devices using, for example, wired or wireless networks. System 295 can also include one or more input devices such as keyboard and mouse, and one or more output devices such as a display and printer.

Computer system 295 can include analysis tools, such as MATLAB from Mathworks of Natick, Massachusetts, that are used to analyze the data. See the

example MATLAB scripts, provided at the Appendix hereto, that can be used to analyze and model the collected data.

Referring now to Figure 12, another system 300 for measuring macromolecular interactions is shown. System 300 is similar to system 200 described above, but includes additional components. Like numbers in system 300 refer to like elements in system 200. System 300 generally includes a dispenser 310, solution reservoirs 222, 224, 326, a mixer 340, a valve 250, a filter 260, a splitter 270, detectors 280, 290, and a computer system 295.

Dispenser 310 includes triple-syringes 212, 214, 316 connected to valves 232, 234, 336 respectively. Solution reservoirs 222, 224, 326 are also connected to three-way valves 232, 234, 336 respectively. In some embodiments, reservoir 222, 224, 326 can each contain a different stock solution. In some embodiments, for example, reservoir 222, 224 each contain a different stock solution and reservoir 326 contains a solvent. Valves 232, 234, 336 are programmable so that valves 232, 234, 336 can be switched between a filling mode and a delivery mode. For example, for the filling mode, valve 232 can be switched to allow syringe 212 to be filled with stock solution from reservoir 222, valve 234 can be switched to allow syringe 214 to be filled with stock solution from reservoir 224, and valve 336 can be switched to allow syringe 316 to be filled with stock solution from reservoir 326. Valves 232, 234, 336 can likewise be switched for the delivery mode to dispense fluid from syringes 212, 214, 316.

A rate of delivery from syringes 212, 214, 316 can be controlled. In one example, the rate of delivery from syringes 212, 214, 316 can be controlled to create a gradient in concentration of the solution that is dispensed by dispenser 310 and/or to control the flow rate to provide a balanced flow rate.

Valves 232, 234, 336 are, in turn, connected to T-junction or mixer 340. Mixer 340 combines, for example, stock solution from syringe 212, stock solution from syringe 214, and stock solution from syringe 316 into a single stream of solution ("solution stream"). In one example, mixer 340 is a stream mixer (Upchurch Scientific, Oak Harbor, WA). Other types of mixers can be used. The solution stream from mixer 340 is delivered to valve 250 as described above.

B. Methods and Uses

The systems and methods of the invention can be used to detect and characterize associations of one or more species of macromolecules in a solution. Examples of macromolecules include polypeptides such as proteins, receptors, antibodies, antibody fragments, monobodies and immunoadhesions, polynucleotides such as DNA and RNA, starches, lipids, cellulose, lignans, macromolecular pharmaceutical compounds, organic or inorganic polymers such as olefins, polyesters, polyethylenes, polyurethanes, and polysaccharides, synthetic rubbers, synthetic lubricants, chitosan, food stabilizers, virus particles, and vaccines. The systems of the invention can detect and characterize self-associations of macromolecules in a homogenous solution. For example, the systems of the invention can detect and characterize macromolecular interactions such as monomer, dimer, or trimer formation, and the like of macromolecules, such as proteins, in a homogenous solution. The systems of the invention can also detect thermal disassociation, denaturation, conformation, and/or purification of macromolecules in a homogenous or heterogeneous solution.

The systems of the invention can also detect macromolecular interactions between one or more macromolecule species in a heterogeneous solution. For example, the systems of the invention can detect and characterize macromolecule interactions between a protein and DNA or RNA, a protein and an antibody, two or more different species of proteins, a protein and an organic polymer, inorganic polymer, or biopolymer, two or more species of organic polymers, two or more species of biopolymers, a biopolymer and an organic polymer, a biopolymer and an inorganic polymer, or a organic polymer and an inorganic polymer. In an embodiment, the systems of the invention can detect macromolecular interactions between a protein and an agonist or antagonist. The agonist or antagonist can be a protein, antibody, antibody fragment monobody, immunoadhesion, or receptor.

In an embodiment, the systems and methods of the invention can be configured for high-throughput analysis of a solution of macromolecules. In such an embodiment, the dispensing module can be connected to a production line for a macromolecular solution and configured to draw a sample of the macromolecular solution from the production line. In such an embodiment, the system can be configured to provide instructions to the production line to maintain, for example, the concentration of solutes or aggregates in the macromolecular solution within

defined production parameters. In an embodiment, the system can be configured with a fraction collector and programmable sample handling robotics. The sample handling robotics can be programmed to sequentially transfer production line samples from the fraction collector to individual pump reservoirs for a series of assays.

The systems and methods of the invention can be used to analyze the safety or efficacy of a formulation comprising a therapeutic agent or macromolecular pharmaceutical compound or therapeutic protein. Aggregation of the pharmaceutical compound or therapeutic protein, for example, can reduce the efficacy of the formulation for treating a disease or disorder. In some instances, aggregation of a pharmaceutical compound or protein results in an aggregate that is toxic. The systems and methods of the invention can also be used to determine how strongly a pharmaceutical compound or protein binds a target molecule. Information related to how strongly a pharmaceutical compound or therapeutic protein binds a target molecule can be used to determine an appropriate dosage for treating a disease or disorder. The systems and methods of the invention can also be used to determine how strongly a pharmaceutical compound or therapeutic protein binds non-target molecule that would reduce the efficacy of the compound or protein.

The systems and methods of the invention can be used for high-throughput analysis of a solution of macromolecules.

Systems 100, 200, and 300 can be used for detection and characterization of reversible associations of one or more species of macromolecules in a solution. For example, dual-syringes 212, 214 of dispenser 210 can introduce a solution of time-varying composition into detectors 280, 290 for simultaneous measurement of laser light scattering at multiple angles and absorbance. Examples of the uses of systems 100, 200, and 300 are provided below.

1. Self-Association

Following a baseline measurement, dispenser 210, under program control as described above, provides a stepwise upward or downward gradient of solute concentration that varies roughly linearly over a period of time. In an embodiment, as little as 1 ml of buffer and 1 ml of a stock solution including at least one macromolecular species with absorbance of 0.1 OD units or greater at the selected

wavelength is sufficient to provide the gradient. The relative intensity of light scattered at multiple angles and concentration of the sample are collected at regular intervals using parallel streams delivered to detectors 280, 290. In an embodiment, the relative intensity of 690-nm light scattered at 90 degrees and the relative absorbance of the solution at 280 nm are recorded as functions of time. Raw data is saved in native ASTRA format in detector 280 and exported as text files to computer system 295 for analysis. The data is analyzed as described below to determine the degree of association/non-association based on the simultaneously generated light scattering and concentration data.

2. Hetero-Association

The process for detecting and quantifying hetero-associations between different macromolecular species, such as two different macromolecular species referred to as A and B, is performed in a manner similar to that described above, with the following modifications. Solution A contains A at w/v concentration w_A^o, and solution B contains B at w/v concentration w_B^o. In an embodiment, solutions A and B are placed in reservoirs 222 and 224, respectively, and loaded into the corresponding syringes 212, 214. A baseline is obtained using a buffer solution from reservoir 258 connected by a peristaltic pump 256 to inlet 252 of valve 250. Next, B is introduced into the scattering/absorbance detectors 280, 290 until a plateau of signal is obtained. Typically, this requires 700 - 800 μl of solution. Next, a temporal gradient of composition B is initiated, during which the fraction of solution B (f_B) in the solution mixture introduced into detectors 280, 290 is decreased and the fraction of solution A ($f_A = 1 - f_B$) is simultaneously increased in stages, over a time period of 5 - 20 minutes, until $f_A = 1$.

Following each stage, corresponding to an increment of 0.05 in the value of f_A, syringe pump 210 pauses for a pre-selected period to allow the solution mixture to equilibrate. Establishment of the complete gradient of only B to only A (or vice versa) requires approximately 1 ml of each solution. In an embodiment, solution A and solution B are at least 95% pure using conventional methods of purification. The entire experiment typically requires 2 ml of each solution, at a concentration of ca. 0.5 mg/ml, or a total of ca. 1 mg of each protein.

C. Data Analysis

1. Self-Association

The data points collected can be analyzed to identify macromolecular interactions. In an embodiment, the data points are collected by system 200 from the simultaneous or substantially simultaneous measurements taken by light scattering detector 280 and concentration detector 290. Many different models are possible. See, for example, Example 1 which describes modeling of self-associating proteins and Example 4 which describes modeling of indefinite self-associating proteins.

In one example, the following process is used to analyze the data collected by system 200. Initially, the absorbance data from detector 290 is converted into time-dependent concentration data using previously measured extinction coefficient(s). In addition, the scattering data from detector 280 is converted to concentration- and angle-dependent values of the Rayleigh ratio for excess (solute) scattering $R(\theta,\{w\})$, where $\{w\}$ denotes the composition of the solution specified by weight/volume concentration of all solute species. Additionally, the value of the optical constant K' is calculated as shown in Equation 1 below:

$$K' = 4\pi^2 n_o^2 \lambda_o^{-4} N_A^{-1}, \quad (1)$$

where n_o denotes the refractive index of buffer, λ_o is the wavelength in vacuum of the scattering light (e.g., 600 nm), and N_A is Avogradro's number. The value of the Rayleigh ratio R at zero scattering angle for a mixture of dilute species can be calculated as follows in Equation 2:

$$R(0,\{w\}) = K' \sum_i \left(\frac{dn}{dw_i}\right)^2 M_i w_i, \quad (2)$$

where w_i denotes the weight/volume concentration and dn/dw_i denotes the specific refractive increment of the ith solute species. If all scattering species have the same chemical composition (e.g., a polymer with a distribution of chain length or a single protein that self-associates to form different oligomeric species), then the refractive

increment of all species is equal, and Equation 2 simplifies to the following Equation 3:

$$R(0, w_{\text{tot}}) = K\sum_i M_i w_i = K w_{tot} M_{\text{W}}, \quad (3)$$

where K approximates $K'(dn/dw)^2$, w_{tot} is the total concentration of solute, and M_{W} is the weight-average molar mass.

The dependence of $R(0,w_{\text{tot}})$ upon w_{tot} is obtained as follows. Data points (typically several thousand) are tabulated as a function of two variables, w_{tot} and $\sin^2(\theta/2)$. A two-dimensional polynomial is created as provided in Equation 4 below:

$$Z(\theta, w_{\text{tot}}) \equiv \frac{R(\theta, W_{\text{tot}})}{K} = \sum_{i=0}^{i_{max}} \sum_{j=1}^{j_{max}} C_{ij} [\sin^2(\theta/2)]^i w_{\text{tot}}^j . \quad (4)$$

The two-dimensional polynomial is fit globally to the entire data set by linear least squares. Combination of Equations 3 and 4 yields the following Equation 5:

$$M_W = \frac{Z(0, w_{\text{tot}})}{w_{\text{tot}}} = \sum_{j=1}^{j_{max}} C_{0j} w_{\text{tot}}^{j-1} . \quad (5)$$

Fisher's F test can be used to determine the minimum values of i_{max} and j_{max} permitting Equation 4 to describe the entire data set to within experimental uncertainty. Globular proteins whose maximum dimension is less than 1/20th of the wavelength of scattering light (ca. 35 nm) typically behave as point particles with no angular dependence of scattering. For these solutes, a lowest acceptable value of $i_{\text{max}} > 0$ is indicative of either an instrumental artifact or the presence of aggregates formed subsequent to pre-filtration of the protein stock solution.

If the data is described to within experimental precision as expressed in Equation 4 with $i_{\text{max}} = 0$ and $j_{\text{max}} = 1$, then M_{W} is independent of concentration over the range of solute concentrations up to that of the stock solution. This result is consistent with one of two possibilities: (1) there exists a single non-self-interacting solute species, with molar mass M equal to M_{W} (= C_{01}); or (2) there exists a mixture

of non-interacting solute species. The presence of multiple solute species can be revealed by size exclusion chromatography, native gel electrophoresis, and/or sedimentation velocity experiments.

If a description of the data according to Equation 4 requires C_{02} to be significantly greater than 0 (i.e., M_W increases with solute concentration), then the present of equilibrium association is indicated. Conversely, if a description of the data according to Equation 4 requires C_{02} to be significantly less than 0 (i.e., M_W decreases with solute concentration), then the solution is exhibiting non-ideal behavior arising from repulsive solute-solute interaction.

When the multi-angle scattering data can be satisfactorily described by Equation 4 with $I_{max} = 0$, indicating a lack of angular dependence, a further simplification is possible. For each time (or concentration) point, the value of Z (= R/K) obtained at 15 scattering angles is averaged, and the results are saved as a table of $\{w_{tot}, <Z>(w_{tot})\}$. This process is referred to as "data condensation." The dependence of $<Z>$ upon w_{tot} is then modeled in the context of a model for equilibrium self-association as indicated below.

For example, for a monomeric protein A in equilibrium with one or more oligomeric species A_i, the molar concentration of each i-mer is given by Equation 6 below:

$$c_i = K_i c_1^i, \tag{6}$$

where c_i denotes the molar concentration of i-mer. Conservation of mass can be expressed as shown in Equation 7 below:

$$c_{iot} = w_{tot} / M_1 = \sum i c_i = \sum i K_i c_1^i . \tag{7}$$

Equation 7 can be solved analytically or numerically for c_1 as a function of w_{tot}, M_1, and the various K_i. Then each of the c_i can be calculated using Equation 6, and Equations 8a and 8b follow:

$$Z = M_W w_{tot} , \text{ where} \tag{8a}$$

$$M_W = \frac{\sum M_i w_i}{\sum w_i} = M_1 \frac{\sum i^2 c_i}{\sum i c_i}. \quad (8b)$$

The values of M_1 and each K_i can be estimated by nonlinear least-squares fitting of Equations 6-8b to the experimentally measured dependence of <Z> upon w_{tot}.

2. Hetero-Association

Many different models are possible for identifying and analyzing hetero-associations in a solution comprising multiple macromolecular species. See, for example, Example 2 which describes the modeling of reversible macromolecular hetero-associations and Example 3 which describes the simultaneous modeling of both self- and hetero-associations in a solution containing multiple species of macromolecules.

The following process can be used to analyze data related to hetero-association of at least two different macromolecule species, A and B. Initially, the time- and angle-dependent Rayleigh ration $R(\theta,t)$ are calculated from the data points.

Data outliers (typically less than 1 percent of the total data points) can then be removed as follows. For each data point collected, the following function shown in Equation 9 is fitted by linear least squares to values or R_i/K obtained from detectors 280 and 290, corresponding to scattering angles $\theta_{(8\text{-}16)}$ between 60 and 142 degrees:

$$R(t)/K = a_0(t) + a_1(t) \sin^2(\theta/2). \quad (9)$$

Next, using the best-fit values of a_0 and a_1, the squared residual corresponding to detectors 4-18 is calculated according to the following Equation 10:

$$\delta_i^2(t) = [R_i(t)/K - a_0(t) - a_1(t) \sin^2(\theta_i/2)]^2. \quad (10)$$

A mean square residual characterizing the data obtained at intermediate scattering angles can then be defined as follows in Equation 11:

$$MSR(t) = \frac{1}{9}\sum_{i=8}^{16}\delta_i^2(t). \quad (11)$$

Data filtering can be accomplished by removing each data point for which $\delta_i^2(t) > 3MSR(t)$. Once data outliers are removed, the Rayleigh ratio can be scaled to a pre-calculated optical constant K defined in a manner similar to that described above (see Equation 1).

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(λ,t) according to Equations 12, 13a, and 13b:

$$f_a(T) = \frac{A(\lambda,T) \quad w_B^o\varepsilon_B(\lambda)}{w_A^o\varepsilon_A(\lambda) \quad w_B^o\varepsilon_B(\lambda)}, \quad (12)$$

$$w_{A,tot}(t) = f_A(t)w_A^o, \quad (13a)$$

$$w_{B,tot}(t) = [1 - f_A(t)]w_B^o, \quad (13b)$$

where $\varepsilon_A(\lambda)$ and $\varepsilon_B(\lambda)$ 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 is no angular dependence of scattering. The processed data is modeled as a two column array of $\{f_A, \langle R\rangle/K\}$, where <R> is the mean value of R obtained from detectors 4-18, with outliers removed as described above. If data exhibits an angular dependence of R, the composition can be modeled using $\lim_{\theta\to 0} R(\theta,\{w\})$, rather than <R>({w}).

Since the solution includes proteins A and B, a variable associated with a particular species A_iB_j bears the subscript ij. Fractional association competence of each protein is denoted by $f_{A,comp}$ and $f_{B,comp}$, respectively. The total molar concentrations of competent protein and incompetent protein (i.e., the certain mass fraction of each protein that is incompetent to form complexes) are given by the following Equations 14a and 14b:

$$c_{X,tot} = f_{X,comp}\, w_{X,tot} / M_X, \tag{14a}$$

$$c_{X,inc} = (1 - f_{X,comp})\, w_{X,tot} / M_X, \tag{14b}$$

where X can be either A or B. An equilibrium association scheme is defined by the specification of one or more equilibrium association constants of the form shown in Equation 15:

$$K_{ij} \equiv \frac{c_{ij}}{c_{10}^i c_{01}^j}, \tag{15}$$

where c_{ij} denotes the molar concentration of A_iB_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 the following Equations 16a and 16b:

$$c_{A,tot} = \sum_{i,j} i c_{ij} = \sum_{i,j} i K_{ij} c_{10}^i c_{01}^j, \tag{16a}$$

$$c_{B,tot} = \sum_{i,j} j c_{ij} = \sum_{i,j} j K_{ij} c_{10}^i c_{01}^j. \tag{16b}$$

For each value of f_A, the corresponding value of <R>/K is calculated as follows:

(1) the values of $w_{A,tot}$ and $w_{B,tot}$ are calculated using Equations (13a) and (13b), 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 $c_{A,tot}$, $c_{A,inc}$, $c_{B,tot}$, and $c_{B,inc}$ are calculated using Equations 14a and 14b;

(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 Equations 16a and 16b;

(4) the values of all c_{ij} are calculated using Equation 15; and

(5) the value of <R>/K is then calculated according to the following Equation 17:

$$< R > / K = M_{10}^2 c_{A,inc} + M_{01}^2 c_{B,inc} + \sum_{i,j} M_{ij}^2 c_{ij} \,. \qquad (17)$$

EXAMPLES

The present invention may be better understood with reference to the following examples. These examples are intended to be representative of specific embodiments of the invention, and are not intended as limiting the scope of the invention.

Example 1

Detection and Characterization of Macromolecular Interactions Between Self-Associating Proteins

The composition dependence of the light scattering of a mixture of macromolecules can be analyzed to yield information about attractive and repulsive interactions between individual species. However, acquisition of such information utilizing conventional batch procedures is a time-consuming and labor-intensive process; hence it is rarely utilized. In this example, we demonstrate a system comprising a liquid dispensing instrument and light scattering and concentration detectors to acquire large quantities of accurate composition-dependent light scattering data rapidly and automatically. Utilizing a novel analytical procedure, the data acquired by the system can be interpreted equally rapidly to yield reliable estimates of the molar mass(es) of macrosolute species and the strength of reversible associations between them.

Materials

Albumin (bovine serum monomer), albumin (chicken egg white), alcohol dehydrogenase (yeast), cytochrome c (horse heart), fibrinogen (bovine plasma, type IV), pepsinogen (porcine stomach), β-lactoglobulin A (bovine milk), b-lactoglobulin B (bovine milk), lysozyme (chicken egg white), and hemoglobin (human) were obtained from Sigma–Aldrich (St. Louis, MO). Chymotrypsinogen A (3· crystallized) was obtained from Worthington Chemical (Freehold, NJ). Except for hemoglobin, all proteins were used without further purification. Hemoglobin was converted to cyanmethemoglobin as described in Benesch et al., 1978, *Biochem. Biophys. Res. Commun.*, 81:1307-1312.

Sample Preparation

Before use, all protein solutions were dialyzed against phosphate-buffered saline (PBS), pH 7.2 (Biosource, Biofluids, USA). Hemoglobin solutions prepared in high-ionic-strength buffers were dialyzed against PBS to which the requisite quantity of NaCl was added. Buffers were prefiltered through Millipore 0.22-lm filters. Protein solutions were prefiltered through 0.02-lm Whatman Anotop filters. Immediately before experiments were performed, buffers and protein solutions were centrifuged at 1000g for 15 min to remove residual particulates and microscopic bubbles. All measurements were carried out at 20 °C.

Instrumentation

Solutions were dispensed by a Hamilton Microlab 900 dual-syringe precision dispenser (Hamilton, Reno, NV) and delivered through a Whatman Anotop 0.1 μm filter to a Wyatt DAWN-EOS multiangle laser light scattering detector (Wyatt Technology, Santa Barbara, CA), equipped with a temperature-regulated K5 flow cell and a Milton Roy SM3100 variable-wavelength UV–visible absorbance detector (Thermo Finnegan, West Palm Beach, FL), installed in parallel as indicated generally and schematically in Figures 1 and 2. The analog output of the absorbance detector (1 V per absorbance unit) was connected to the AUX1 input of the DAWN-EOS, and data from the scattering and absorbance detectors were collected simultaneously using ASTRA software (Wyatt Technology; Release 4.90.04). Adjustment of the flow rate in each of the parallel flow paths was necessary to ensure that absorbance measured at a particular time point corresponds to the composition of solution scattering light at the same time point.

Refractive increments of proteins were measured using a thermostatted Leica ARIAS 500 Abbe′ refractometer (Reichert Instruments, Buffalo, NY) and corrected for differences between the measurement wavelengths of the refractometer (589 nm) and the DAWN-EOS (690 nm) according to Perlmann and Longsworth, 1948, *J. Am. Chem. Soc.*, 70:2719-2724. Extinction coefficients of proteins at the appropriate wavelengths were measured by injection of protein solutions of known concentration into the absorbance detector, and applicability of Beers' Law was confirmed for all proteins examined.

Data Analysis

All calculations were performed automatically using scripts and functions, written and executed in MATLAB as provided in the Appendix hereto (Mathworks, Natick, MA). Using previously measured extinction coefficient(s), absorbance data was converted into time-dependent concentration data and, following procedures provided in the DAWN-EOS instruction manual, the raw scattering data was converted to concentration- and angle-dependent values of the Rayleigh ratio for excess (solute) scattering $R(\theta,\{w\})$, where $\{w\}$ denotes the composition of the solution specified by weight/volume concentration of all solute species. Additionally, the value of the optical constant K' was calculated as shown in Equation 1 below:

$$K' = 4\pi^2 n_o^2 \lambda_o^{-4} N_A^{-1}, \quad (1)$$

where n_o denotes the refractive index of buffer, λ_o is the wavelength in vacuum of the scattering light (e.g., 600 nm), and N_A is Avogradro's number. The value of the Rayleigh ratio R at zero scattering angle for a mixture of dilute species can be calculated as follows in Equation 2:

$$R(0,\{w\}) = K' \sum_i \left(\frac{dn}{dw_i}\right)^2 M_i w_i, \quad (2)$$

where w_i denotes the weight/volume concentration and dn/dw_i denotes the specific refractive increment of the ith solute species. If all scattering species have the same chemical composition (e.g., a polymer with a distribution of chain length or a single protein that self-associates to form different oligomeric species), then the refractive increment of all species is equal, and Equation 2 simplifies to the following Equation 3:

$$R(0,w_{tot}) = K \sum_i M_i w_i = K w_{tot} M_W, \quad (3)$$

where K approximates $K'(dn/dw)^2$, w_{tot} is the total concentration of solute, and M_W is the weight-average molar mass.

The dependence of $R(0,w_{tot})$ upon w_{tot} was obtained as follows. Data points (typically several thousand) were tabulated as a function of two variables, w_{tot} and $\sin^2(\theta/2)$. A two-dimensional polynomial was created as provided in Equation 4 below:

$$Z(\theta, w_{tot}) \equiv \frac{R(\theta, W_{tot})}{K} = \sum_{i=0}^{imax} \sum_{j=1}^{jmax} C_{ij} [\sin^2(\theta/2)]^i w_{tot}^j . \quad (4)$$

The two-dimensional polynomial was fit globally to the entire data set by linear least squares. Combination of Equations 3 and 4 yielded the following Equation 5:

$$M_W = \frac{Z(0, w_{tot})}{w_{tot}} = \sum_{j=1}^{jmax} C_{0j} w_{tot}^{j-1} . \quad (5)$$

Fisher's F test was used to determine the minimum values of i_{max} and j_{max} permitting Equation 4 to describe the entire data set to within experimental uncertainty. Globular proteins whose maximum dimension was less than 1/20th of the wavelength of scattering light (ca. 35 nm) typically behaved as point particles with no angular dependence of scattering. For these solutes, a lowest acceptable value of $i_{max} > 0$ was indicative of either an instrumental artifact or the presence of aggregates formed subsequent to pre-filtration of the protein stock solution.

Results

A. Characterization of equilibrium self-association

If the data are described within experimental precision as expressed in Equation 4 with $i_{max} = 0$ and $j_{max} = 1$, then M_W is independent of concentration over the range of solute concentrations up to that of the stock solution. This result is consistent with one of two possibilities: (1) there exists a single non-self-interacting solute species, with molar mass M equal to M_W (= C_{01}); or (2) there exists a mixture of non-interacting solute species. The presence of multiple solute species can be

revealed by such techniques as size exclusion chromatography, native gel electrophoresis, and/or sedimentation velocity experiments.

If a satisfactory description of the data according to Equation 4 requires C02 to be significantly greater than 0 (i.e., MW increasing with solute concentration), then the presence of equilibrium association is indicated. Conversely, if a satisfactory description of the data according to Equation 4 requires C02 to be significantly less than 0 (i.e., MW decreasing with solute concentration), then the solution is exhibiting non-ideal behavior arising from repulsive solute–solute interaction [K.A. Stacey, 1956, *Light-Scattering in Physical Chemistry*, Academic Press, New York]. We have considered only self-association of a single ideal solute component.

When the multi-angle scattering data can be satisfactorily described by Equation 4 with $I_{max} = 0$, indicating a lack of angular dependence, a further simplification is possible. For each time (or concentration) point, the value of Z (= R/K) obtained at 15 scattering angles is averaged, and the results are saved as a table of $\{w_{tot}, \langle Z \rangle (w_{tot})\}$. This process is referred to as "data condensation." The dependence of $\langle Z \rangle$ upon w_{tot} is then modeled in the context of a model for equilibrium self-association as indicated below.

For example, for a monomeric protein A in equilibrium with one or more oligomeric species A_i, the molar concentration of each i-mer is given by Equation 6 below:

$$c_i = K_i c_1^i, \tag{6}$$

where c_i denotes the molar concentration of i-mer. Conservation of mass can be expressed as shown in Equation 7 below:

$$c_{iot} = w_{tot} / M_1 = \sum i c_i = \sum i K_i c_1^i . \tag{7}$$

Equation 7 can be solved analytically or numerically for c_1 as a function of w_{tot}, M_1, and the various K_i. Then each of the c_i can be calculated using Equation 6, and Equations 8a and 8b follow:

$$Z = M_{\mathrm{W}} w_{\mathrm{tot}}, \text{ where} \tag{8a}$$

$$M_{\mathrm{W}} = \frac{\sum M_i w_i}{\sum w_i} = M_1 \frac{\sum i^2 c_i}{\sum i c_i}. \tag{8b}$$

The values of M_1 and each K_i can be estimated by nonlinear least-squares fitting of Equations 6-8b to the experimentally measured dependence of $\langle Z \rangle$ upon w_{tot}.

B. Nonassociating proteins

The results of analysis of data obtained for ovalbumin is plotted in Figure 3. Ovalbumin is a known non-associating protein. Figure 3 includes a raw data set of 16,755 data points, and shows scaled 90 degree light scattering (690 nm) and absorbance (280 nm) data plotted as a function of elapsed time through a dilution gradient. The two curves are nearly superimposed on one another.

Figures 4A, 4B, 5A, and 5B show further refinement of the data shown in Figure 3. For example, Z values are plotted as a function of w_{tot} and $\sin^2(\theta/2)$ in Figure 4A. Also plotted is the best fit of Equation 4 with $i_{\max} = 0, j_{\max} = 1$, and $C_{01} = 44{,}388 \pm 22$ (95 percent confidence limits). The corresponding best-fit residuals are plotted in Figure 4B. Figures 5A and 5B show the results of the same analysis applied to a filtered subset of the initial data set, obtained by deleting all of the points in the original data set with values of the squared best fit residual greater than three times the value of the mean squared best-fit residual. This subset has 16,696 data points, and the best-fit value of $C_{01} = 44{,}341 \pm 8$. It can be seen that that filtering procedure does not significantly alter the result of the analysis, indicating that the small number of outliers in the raw data set have no significant effect on the determination of the molar mass.

In Figure 6, the value of log M determined for several proteins by the method described above is plotted against the value of log M for the corresponding protein obtained from literature. The dashed line indicates equal-valued x and y coordinates. Proteins shown include: (1) fibrinogen; (2) alcohol dehydrogenase; (3) bovine serum albumin (nonequilibrium mixture of monomer + oligomers); (4) hemoglobin; (5) bovine serum albumin; (6) ovalbumin; (7) pepsinogen; (8) β-

lactoglobulin (mixture of A and B); (9) β-lactoglobulin A; (10) chymotrypsinogen A; (11) lysozyme; and (12) cytochrome *c*.

C. Self-associating proteins

Some of the proteins examined exhibited an improvement in the quality of fit of Equation, as measured by the magnitude of the sum of squared residuals, when j_{max} was increased from 1 to 2, and best-fit values of C_{02} were found to be significantly greater than zero. A further increase of j_{max} from 2 to 3 did not result in further significant lowering of the sum of squared residuals. These data sets were then condensed as described above. Ideal monomer–dimer self-association models based upon Equations 6-8 were fit to the condensed data sets by the method of nonlinear least squares to obtain best-fit estimates of M_1 and log K_2 (M^{-1}).

Experimentally measured values of $<R>/K$ for β -lactoglobulin are plotted as a function of w_{tot} in Figure 7. The data points shown represent the experimental data, and the line illustrates the best fit of a monomer–dimer equilibrium association model calculated with best-fit parameter values shown in Table 1 below. The dashed line represents the hypothetical dependence of $<R>/K$ upon w_{tot} in the absence of self-association, calculated using best-fit monomer molecular weight. Best-fit parameter values for this and other self-associating proteins are summarized in Table 1 below, which summarizes results of modeling light scattering data in the context of a model for ideal monomer–dimer equilibrium self-association.

Table 1

Protein	Number of data sets	M_1 (best-fit)	log K_2 (M^{-1}) (best-fit)
β-Lactoglobulin	3	19,600 [18,800; 20,500]	5.3 [5.0; 5.7]
Chymotrypsinogen	2	23,600 [23,100; 24,000]	3.25 [3.1; 3.35]
Cyanmethemoglobin [NaCl] = 0.15 M	1	35,700[a] [33,900; 38,000]	5.8 [5.2; 7.1]
Cyanmethemoglobin [NaCl] = 1 M	2	32,200[a] [29,500; 37,000]	4.7 [4.1; 5.2]
Cyanmethemoglobin [NaCl] = 2 M	2	33,400[a] [29,500; 37,000]	4.1 [3.8; 4.5]

Bracketed values following best-fit values represent lower and upper 95% confidence limits of estimate.

[a] The association process characterized corresponds to $2(\alpha\beta) \rightleftharpoons \alpha_2\beta_2$.

Discussion

Measurement and analysis of excess static light scattering of macromolecules in solution has traditionally been carried out in batch mode. A series of solutions containing a macrosolute at different concentrations is prepared and then, in sequence, each solution is introduced into the scattering cell and the scattering is measured at multiple angles. For each solute concentration, an apparent weight-average molar mass and, for sufficiently large macrosolutes, the z-average radius of gyration of the solute are determined by linear regression of $R(\theta, w_{tot})/Kw_{tot}$ or $Kw_{tot}/R(\theta, w_{tot})$ [B.H. Zimm, 1948, *J. Chem. Phys.*, 16:1099-1116; K.A. Stacey, 1956, *Light-Scattering in Physical Chemistry*, Academic Press, New York]. The presence of significant solute–solute interaction under a particular set of experimental conditions is manifested as a concentration dependence of the apparent weight-average molar mass [Tojo et al., 1985, *J. Biol. Chem.*, 260:12607-12614; Osborne and Steiner, 1972, *Arch. Biochem. Biophys.*, 152:849-855; Yamaguchi and Adachi, 2002, *Biochem. Biophys. Res. Commun.*, 290:1382-1387].

The system and methods described herein improve on the traditional approach described above in several respects. The total amount of macrosolute required for the complete analysis is many times smaller. A more complete characterization of concentration dependence of solution properties is achieved via a continuous gradient of concentration in contrast to a few discrete concentrations. The process of data acquisition is much more rapid. Attempts to model values of $R(\theta, w_{tot})/Kw_{tot}$ or $Kw_{tot}/R(\theta, w_{tot})$, which are themselves obtained by regression, as functions of the independent variable wtot are problematic both statistically and numerically. These variables, which are conventionally treated in the context of regression as nominally dependent variables, are not dependent variables but rather extremely complex compound variables. Moreover, the precision of either variable diverges in the sought limit of low concentration. In contrast, in the present analysis, the concentration dependence of zero-angle scattering is obtained via robust and essentially instantaneous global modeling of $R(\theta, w_{tot})/K$ at all scattering angles and all concentrations, with no loss of precision beyond that inherent in the signal/noise ratio of the raw data at low concentrations.

Previous investigations of the behavior of β-lactoglobulin [Kelly and Reithel, 1971, *Biochemistry*, 10:2639-2644] and chymotrypsinogen A [Osborne and Steiner, 1972, *Arch. Biochem. Biophys.*, 152:849-855] have established that these proteins do self-associate under conditions comparable to those of the present experiments, but a quantitative comparison between earlier and present results is not possible due to significant differences in temperature, ionic strength, and/or buffer composition. However, equilibrium constants for dimer–tetramer association of oxy- and carboxyhemoglobin, measured previously under conditions almost identical to those employed here, have been tabulated [Antonini and Brunori, 1971, *Hemoglobin and Myoglobin in their Reactions with Ligands*, North Holland, Amsterdam]. Since cyanmet-, oxy-, and carboxyhemoglobin share the same quaternary structure [Marden et al., 1991, *Biophys. J.*, 60:770-776]. The equilibrium association constant of cyanmethemoglobin obtained in the present work agrees with the tabulated values for oxy-and carboxyhemoglobin to within experimental uncertainty at all three values of the ionic strength.

The analysis of equilibrium self-association presented herein is based upon the assumption that equilibration between the various associating species is rapid with respect to the rate of change of composition of the solution. To check the validity of this assumption, a delay time was introduced following the addition of successive increments of solution, and the scattering versus time curve was examined for the appearance of relaxations that are significantly slower than the mixing time as monitored by the rate of change of absorbance. In this manner we observed a significant scattering lag accompanying the addition of large quantities of buffer to small quantities of stock β-lactoglobulin, which was attributed to the time required for protein association to reequilibrate following rapid dilution. A corresponding lag was not observed in the downward gradient of concentration, where dilution is gradual rather than abrupt. We therefore subsequently analyzed descending gradients of concentration in the self-associating protein systems and, for each protein studied, performed experiments at different rates of concentration change to ascertain that the derived dependence of $\langle R \rangle / K$ upon total concentration was independent of this rate.

There is an analogy, both thermodynamic and methodological, between the present approach to measurement and analysis of light scattering and the mea-

surement and analysis of sedimentation equilibrium. In the light scattering experiment, one simultaneously measures concentration and total solute scattering, a property that depends upon the product of solute mass and refractive increment of each solute species and upon the interactions, both attractive and repulsive, between solute molecules [K.A. Stacey, 1956, *Light-Scattering in Physical Chemistry*, Academic Press, New York; W.H. Stockmayer, 1950, J. Chem. Phys. 18:58-61]. In the centrifugation experiment, one simultaneously measures concentration and the equilibrium gradient of solute(s), a property that depends upon the product of solute mass and density increment of each solute species and upon the interactions, both attractive and repulsive, between solute molecules [Winzor et al., 1998, *Biochemistry*, 37:2226-2233; Wills et al., 2000, *Biophys. J.*, 79:2178-2187; Zorrilla et al., 2004, *Biophys. Chem.*, 108:89-100]. In both types of experiment, information about solute–solute interactions is obtained from observed differences between the measured property of the solution and the expected sum of the properties of isolated (noninteracting) solutes. The systems and methods of the invention acquire information about reversible associations that is comparable in scope and resolution to that currently obtainable from sedimentation equilibrium. Since the systems and methods of the invention permit extremely rapid acquisition and analysis of composition-dependent data, it may be used to characterize reversible associations evolving with time and at equilibrium.

Example 2

Detection and Characterization of Reversible Macromolecular Hetero-Associations

Quantitative characterization of reversible macromolecular associations between different species of biological macromolecules in solution assists in the understanding of how complex biochemical systems respond to changes in composition and environmental variables. In this example, we demonstrate a system comprising a liquid dispensing instrument and light scattering and concentration detectors can rapidly detect and characterize heteroassociations between two different protein species in solution.

Materials

All proteins were obtained from Sigma-Aldrich (St. Louis, MO), dialyzed against the appropriate buffers as described below, and used without further purification. Concentrations were determined from the absorbance at 280 nm using the following standard values for absorbance in OD units per cm path length for a 1 g/l solution: BSA, 0.65; fibrinogen, 1.20; chymotrypsin, 2.04; soybean trypsin inhibitor, 0.94. Refractive increments were determined as described in Example 1, and found to be equal to 0.185 ± 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 in Example 1. Measurements of light scattering were carried out at 20 °C.

Instrumentation

Solutions were dispensed by a Hamilton Microlab 900 dual-syringe precision dispenser (Hamilton, Reno, NV) and delivered through a Whatman Anotop 0.1-lm filter to a Wyatt DAWN-EOS multiangle laser light scattering detector (Wyatt Technology, Santa Barbara, CA), equipped with a temperature-regulated K5 flow cell and a Milton Roy SM3100 variable-wavelength UV–visible absorbance detector (Thermo Finnegan, West Palm Beach, FL), installed in parallel as shown schematically in Figure 2 (system 200). 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 Procedure

Dilution experiments conducted to detect and quantify self-association of a single protein were performed as described in Example 1. The protocol described in Example 1 was modified as follows to detect and quantify association between two different macromolecular solutes, referred to as A and B. Solution A contains A at w/v concentration $w_A°$ and solution B contains B at w/v concentration $w_B°$. solutions A and B were placed in reservoirs 222 and 224 respectively. A baseline was obtained using buffer loaded into reservoir 258 then pure B is introduced into light scattering detector 280 and absorbance detector 290 until a plateau of signal is obtained. Typically this requires 700 - 800 μl of solution. Then a temporal gradient

of composition was initiated, during which the fraction of solution B (f_B) in the solution mixture introduced into detectors 280 and 290 is decreased and the fraction of solution A ($f_A = 1-f_B$) is simultaneously increased in stages, over a time period of 5 - 20 minutes, 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. Establishment of the complete gradient of only B to only 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.

Data Analysis

All calculations were performed automatically using scripts and functions written and executed in MATLAB (Mathworks, Natick, MA). The time- and angle-dependent Rayleigh ration $R(\theta,t)$ were calculated from the unprocessed scattering data as described in Appendix A of the ASTRA for Windows User's Guide (Wyatt Technology, Santa Barbara, CA). Data outliers (typically less than 1 percent of the total data points) were removed as follows. For each data point collected, the following function shown in Equation 9 was fitted by linear least squares to values or R_i/K obtained from detectors 280 and 290, corresponding to scattering angles $\theta_{(8-16)}$ between 60 and 142 degrees:

$$R(t)/K = a_0(t) + a_1(t)\sin^2(\theta/2). \quad (9)$$

Next, using the best-fit values of a_0 and a_1, the squared residual corresponding to detectors 4-18 was calculated according to the following Equation 10:

$$\delta_i^2(t) = [R_i(t)/K - a_0(t) - a_1(t)\sin^2(\theta_i/2)]^2. \quad (10)$$

A mean square residual characterizing the data obtained at intermediate scattering angles was then be defined as follows in Equation 11:

$$MSR(t) = \frac{1}{9}\sum_{i=8}^{16}\delta_i^2(t). \quad (11)$$

Data filtering was accomplished by removing each data point for which $\delta_i^2(t) > 3MSR(t)$. Once data outliers are removed, the Rayleigh ratio was scaled to a pre-calculated optical constant K defined in a manner similar to that described above (see Equation 1).

The fraction of solution A and the time-dependent w/v concentrations of A and B were calculated from the wavelength- and time-dependent absorbance A(λ,t) according to Equations 12, 13a, and 13b:

$$f_a(T) = \frac{A(\lambda,T) \quad w_B^o \varepsilon_B(\lambda)}{w_A^o \varepsilon_A(\lambda) \quad w_B^o \varepsilon_B(\lambda)}, \tag{12}$$

$$w_{A,tot}(t) = f_A(t) w_A^o, \tag{13a}$$

$$w_{B,tot}(t) = [1 - f_A(t)] w_B^o, \tag{13b}$$

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

When all solute species were small relative to the wavelength of scattering light (i.e., maximum dimension less than ca. 40 nm), there was no angular dependence of scattering. The processed data was modeled as a two column array of $\{f_A, \langle R \rangle / K\}$, where <R> is the mean value of R obtained from detectors 4-18, with outliers removed as described above. If data exhibited an angular dependence of R, the composition was modeled using $\lim_{\theta \to 0} R(\theta, \{w\})$, rather than <R>({w}).

Since the solution included proteins A and B, a variable associated with a particular species A_iB_j is denoted with the subscript ij. Fractional association competence of each protein was denoted by $f_{A,comp}$ and $f_{B,comp}$, respectively. The total molar concentrations of competent protein and incompetent protein (i.e., the certain mass fraction of each protein that is incompetent to form complexes) were calculated by the following Equations 14a and 14b:

$$c_{X,tot} = f_{X,comp} \, w_{X,tot} / M_X, \tag{14a}$$

$$c_{X,inc} = (1 - f_{X,comp})\, w_{X,tot} / M_X, \qquad (14b)$$

where X can be either A or B. An equilibrium association scheme was defined by the specification of one or more equilibrium association constants of the form shown in Equation 15:

$$K_{ij} \equiv \frac{c_{ij}}{c_{10}^i c_{01}^j}, \qquad (15)$$

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

Conservation of mass was expressed by the following Equations 16a and 16b:

$$c_{A,tot} = \sum_{i,j} i c_{ij} = \sum_{i,j} i K_{ij} c_{10}^i c_{01}^j, \qquad (16a)$$

$$c_{B,tot} = \sum_{i,j} j c_{ij} = \sum_{i,j} j K_{ij} c_{10}^i c_{01}^j. \qquad (16b)$$

For each value of f_A, the corresponding value of <R>/K was calculated as follows:

(1) the values of $w_{A,tot}$ and $w_{B,tot}$ were calculated using Equations (13a) and (13b), 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 $c_{A,tot}$, $c_{A,inc}$, $c_{B,tot}$, and $c_{B,inc}$ were calculated using Equations 14a and 14b;

(3) given test values of M_{10}, M_{01}, and the log K_{ij}, the values of c_{10} and c_{01} were obtained by either analytical or numerical solution of Equations 16a and 16b;

(4) the values of all c_{ij} were calculated using Equation 15; and

(5) the value of <R>/K was calculated according to the following Equation 17:

$$< R > / K = M_{10}^2 c_{A,inc} + M_{01}^2 c_{B,inc} + \sum_{i,j} M_{ij}^2 c_{ij}. \qquad (17)$$

Best-fit values of two solute component model parameters were determined by non-linear least-squares fitting of a model of the type described above or in AM to the appropriate data set(s) [Press et al., 1987, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, Cambridge]. 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). 95% confidence limits of best-fit model parameters were determined via sum-of-squares profiling [Saroff, H. A., 1989, *Analytical Biochem.*, 176:161-169] combined with the Fisher F-test for equality of variances [Press et al., 1987, *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, Cambridge].

Results

Fibrinogen + bovine serum albumin (BSA). In order to test the accuracy and precision of our measurements and the correctness of our calculations, control experiments were carried out on mixtures of two proteins that were known not to self- or hetero-associate under the conditions of the experiment [Fernandez and Minton, 1999, *Biochemistry*, 38:9379-9388]. Measurements were made on solutions in phosphate-buffered saline (0.05 M Na/K-phosphate + 0.15 M NaCl, pH 7.2). Figures 8A, 8B, and 8C show the values of <R>/K obtained from five dilution experiments, carried out as described in Appendix 3, plotted against $w_{A,tot}$ and $_{B,tot}$. Also plotted is the best fit of the function:

$$R/K = M_A w_{A,tot} + M_B w_{B,tot} \qquad (18)$$

which is the special case of Equation 2 in the absence of self- or hetero-association.

As seen in Figures 8A-8C, the data obtained from dilution experiments carried out on all five solution mixtures can be accounted for by Equation 18. It follows that, in the absence of self- and hetero-association, scattering should depend linearly upon f_A:

$$R/K = M_B w_B^\circ + (M_A w_A^\circ - M_B w_B^\circ) f_A \qquad (19)$$

Thus 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) was a qualitative indicator of interactions between two macrosolutes in solution. 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). Plots of absorbance and scattering data collected by system 200 from a solution A of chymotrypsin (2.04) and a solution B of trypsin inhibitor (0.94) are shown in Figures 9A and 9B, respectively. In Figure 10, the scaled Rayleigh ratio $R(\theta,t)/K$ calculated from the data collected in the experiment shown in Figures 9A and 9B is plotted against $\sin^2(\theta/2)$ and $A_{280}(t)$, thereby eliminating time as an explicit variable. In Figure 11, the value of $\langle R \rangle/K$, calculated from the data plotted in Figures 9A and 9B, is plotted as a function of f_A. Also plotted are the calculated best fits of three models. Model 1 postulates only a simple 1-1 association $A + B \rightleftharpoons AB$ (dotted line). Models 2 and 3 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 $A + B \rightleftharpoons AB$ K_{11} (where $K_{11} = 2K_{site}$) and $2A + B \rightleftharpoons A_2B$ K_{21} (where $K_{21} = K^2_{site}$), where K_{site} is an intrinsic equilibrium association constant for binding to either individual site. In model 2 (dashed line), the fractional association competence of chymotrypsin ($f_{A,comp}$) is constrained to be unity, while in model 3 (solid line), $f_{A,comp}$ is allowed to vary to achieve a best fit. Upper panel of Figure 11 shows best fit residuals of the constrained (dashed curve) and relaxed (solid curve) models.

The present determination of stoichiometry and affinity of the interaction between chymotrypsin and soybean trypsin inhibitor agrees well with a previous characterization of this interaction obtained under very similar experimental conditions via analysis of sedimentation equilibrium [Quast and Steffen, 1975, *Hoppe-Seyler's Z. Phys. Chem.*, 356:617-620]. 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_B obtained using model 3 is in significantly better agreement with published values of the molar mass of soybean trypsin inhibitor [Wu and Scheraga, 1962, *Biochemistry*, 1:698-705]. More importantly, it should be noted that the results plotted in Figure 11 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.

Example 3

Simultaneous Detection and Quantitative Characterization of both Self- and Hetero-Association Equilibria in a Solution Containing Two Protein Components

In this example, we demonstrate a system comprising a liquid dispensing instrument and light scattering and concentration detectors can rapidly detect and characterize pH-dependent reversible heteroassociation between two different protein species and self-association of one of the protein species taking place in the same solution under acidic conditions.

Materials

Chymotrypsin and bovine pancreatic trypsin inhibitor (BPTI) were obtained from Sigma (St. Louis, MO), dialyzed against phosphate buffer, 0.05 M Na Phosphate + 0.2 M NaCl, previously titrated to the indicated pH value, 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/cm pathlength for a 1 g/l solution: chymotrypsin, 2.04; bovine pancreatic trypsin inhibitor, 0.658. Refractive increments were determined as described in Example 1 and found to be equal to 0.185 ± 0.003 ml/g at 20°C for chymotrypsin and BPTI. Immediately before light scattering measurement, solutions were prefiltered and centrifuged as described in Example 1. Measurements of light scattering were carried out at 20°C.

Instrumentation

Experiments were conducted utilizing the instrumentation described in Example 2. A schematic of the instrumentation is shown in Figure 2.

Experimental Procedure

Experiments were conducted as described in Examples 1 and 2. In experiments on mixtures of two different macromolecular solutes (referred to as A and B), a solution of A was loaded into one reservoir (222) and a solution of B was loaded into the second reservoir (224). By simultaneously increasing the flow rate from one syringe and decreasing the flow rate from the second syringe, a composition gradient is established in which the concentration of one protein increases and the concentration of the second protein decreases. The mole fraction of one protein in the mixture gradually increases from 0 to 1, as the mole fraction of the second protein decreases from 1 to 0.

Data Analysis

Raw data obtained from the experiments was processed as described in Example 2 and saved as files of $\{f_A,\langle R\rangle/K\}$, where $\langle R\rangle/K$ denotes the Rayleigh ratio averaged from data obtained by multiple detectors scaled to an optical constant K as defined in Examples 1 and 2, and f_A denotes the faction of the solution containing component A in the two-component solution mixture. All calculations were performed automatically using scripts and functions written and executed in MATLAB (Mathworks, Natick, MA).

Absorbance data were converted to values of $w_{A,tot}$ and $w_{B,tot}$ as described in Example 2. To calculate the value of R/K as a function of $w_{A,tot}$ and $w_{B,tot}$, a model specifying the equilibrium concentrations of all macromolecular solute species present in detectable quantity as a function of the total w/v concentration of each protein component was constructed. The general procedure for constructing the model is described in Example 2.

On the basis of known association properties of chymotrypsin and BPTI (Aune et al., 1971, *Biochem.*, 10:1609-1617; Vincent et al., 1973, *Eur. J. Biochem.*, 38:365-372), the following species were postulated: monomeric A, monomeric B, a heterodimer (AB), and a homodimer (AA). A property of species A_iB_j is indicated by the subscript $\{ij\}$. We defined the following equilibrium constants for heteroassociation of A and B, and for self-association of A, respectively:

$$K_{11} = \frac{c_{11}}{c_{10}c_{01}}, \qquad (20)$$

$$K_{20} = \frac{c_{20}}{c_{10}^2}. \qquad (21)$$

The equations of conservation of mass are then

$$c_{A,tot} = \frac{w_{A,tot}}{M_A} = c_{10} + 2c_{20} + c_{11} = c_{10} + 2K_{20}c_{10}^2 + K_{11}c_{10}c_{11}, \qquad (22)$$

$$c_{B,tot} = \frac{w_{B,tot}}{M_B} = c_{01} + c_{11} = c_{01} + K_{11}c_{10}c_{01}. \qquad (23)$$

Given the experimentally determined values of $w_{A,tot}$, $w_{B,tot}$, and test values of M_A, M_B, K_{11}, and K_{20}, Eqs. 22 and 23 may be solved numerically for the values of c_{10} and c_{01}, and c_{11} and c_{20} are then obtained via Eqs. 20 and 21. The scaled Rayleigh ratio is then calculated as described in Example 2

$$R/K = M_{10}^2 c_{10} + M_{01}^2 c_{01} + M_{11}^2 c_{11} + M_{20}^2 c_{20}. \qquad (24)$$

Results

Experiments on mixtures of chymotrypsin and BPTI were carried out over a range of pH values. The dependence of $<R>/K$ on f_A obtained at three pH values (pH 4.4, pH 5.4, and pH 8.0) is shown in Figs. 13A-13C. Initially, unsuccessful attempts were made to analyze the composition gradient data in the context of a simple 1-1 hetero-association model. It was soon realized that derived values of the molar mass of chymotrypsin were dependent upon pH and unrealistically high at low pH values. Reference to the literature revealed that chymotrypsin is known to dimerize significantly under acid conditions (Aune et al., 1971, *Biochem.*, 10:1609-1617; Aune et al., 1971, *Biochem.*, 10-1617-1622). Subsequently, the data were analyzed in the context of the model described by Eqs. 20-24. This is a further example of how the results obtained using the present technique can rapidly guide an investigator to the correct choice of association model.

To obtain the maximum amount of information about this two-component system, three separate experiments were carried out on solutions prepared at each of

several different pH values (pH 4.4, pH 5.4, and pH 8.0). Dilution (one-component) experiments of the type described in Example 1 were carried out on solutions of each of the two proteins. The composition gradient (two-component) experiment described in Example 2 was then carried out on a time-varying mixture of the two proteins.

Two alternate modeling procedures were used. In the first procedure (single-scan modeling), only the results of the two-component experiment were fit by the model equations. The best fit of the hetero- plus self-association model obtained in this manner at three pH values (pH 4.4, pH 5.4, and pH 8.0) was plotted together with the data in Figs. 13A-13C. Using the values of best-fit parameters, the concentrations of individual species and the contributions of individual species to the total scattering at pH 5.4 were calculated and plotted in Figs. 14A and 14B. Each line in Fig. 14A and 14B represents an individual species: chymotrypsin (A), BPTI (B), hetero-association of chymotrypsin and BPTI (AB), self-association of chymotrypsin (A_2). Fig. 14A shows the calculated concentrations of the individual species. Fig. 14B shows the calculated concentrations of individual species (A, B, and AB) to the total scattering profile (total).

In the second procedure (global modeling), a compound model was constructed to simultaneously fit the results of the two-component experiment together with the results from either one or two one-component experiments, to simultaneously evaluate best-fit values of M_A, M_B, $\log K_{20}$, and $\log K_{11}$. Best-fit values and uncertainties of equilibrium association constants determined by single scan and global modeling are shown in Table 2. In Table 2, values in bold font were obtained by modeling the result of a single composition gradient experiment. Values in non-bold font were obtained by modeling the results of the composition gradient experiment together with the results of a dilution experiment carried out on a solution of chymotrypsin alone, and in some cases, with the results of a dilution experiment carried out on a solution of BPTI alone.

Table 2

pH	log K_{20}	log K_{11}	$M_A/10^3$	$M_B/10^3$
4.4	**4.0 (±0.25)**	**4.5 (±0.45)**	**27.0 (±0.35)**	**5.3 (±0.35)**
	4.3 (±0.2)	5.0 (–0.4, + 0.6)	24.7 (±0.15)	5.6 (±0.3)
4.9	**3.8 (– ∞, + 0.5)**	**5.4 (±1.3)**	**25.4 (±0.25)**	**5.4 (±0.7)**
	4.1 (±0.15)	5.85 (–0.5, + 0.6)	23.9 (±0.1)	5.7 (±0.2)
5.4	**3.7 (±0.2)**	**6.2 (–0.6, + 1.0)**	**23.6 (±0.1)**	**6.1 (±0.25)**
	3.5 (±0.2)	6.0 (±0.5)	24.3 (±0.8)	5.9 (±0.2)
6.7	**3.2 (–0.7, + 0.2)**	**6.8 (–0.7, + ∞)**	**22.9 (±0.1)**	**5.9 (±0.3)**
	2.8 (– ∞, + 0.3)	6.4 (–0.4, + 1.0)	23.5 (±0.1)	5.7 (±0.2)
7.3	**3.2 (±0.2)**	**7.0 (–0.6, + ∞)**	**23.1 (±0.05)**	**5.9 (±0.2)**
	3.0 (–0.5, + 0.3)	7.1 (–0.7, + ∞)	23.2 (±0.5)	5.9 (±0.35)
8.0	**3.1 (– ∞, + 0.4)**	**7.2 (–0.9, + ∞)**	**21.2 (±0.05)**	**5.2 (±0.2)**
	2.7 (– ∞, + 0.3)	6.8 (–0.5, + ∞)	21.4 (±0.3)	5.2 (±0.1)

Values in bold were obtained by modeling the result of a single composition gradient experiment. Values in non-bold font were obtained by modeling the results of the composition gradient experiment together with the results of a dilution experiment carried out on a solution of chymotrypsin alone and, in some cases, with the results of a dilution experiment carried out on a solution of BPTI alone.

Best-fit values of equilibrium constants for hetero- and self-association were plotted as functions of pH in Figs. 15A and 15B. Fig. 15A shows the best-fit values of equilibrium constants for hetero-association of chymotrypsin and BPTI plotted as a function of pH. In Fig. 15A, circles represent values obtained by modeling a single composition gradient experiment, squares represent values obtained by global modeling of a composition gradient together with one or two dilution experiments conducted on individual proteins, diamonds represent values reported by Vincent and Lazdunksi (1973, *Eur. J. Biochem.*, 38:365-372) calculated from the ratio of measured association and measured rate constants, and triangles represent values reported by Rigbi as quoted in Vincent and Lazdunski. Fig. 15B shows the best-fit values of equilibrium constants for self-association of chymotrypsin plotted as a function of pH. In Fig. 15B, triangles represent values obtained by modeling a singles composition gradient experiment, squares represent values obtained from modeling a single dilution experiment conducted on pure chymotrypsin, circles represent values obtained by global modeling of composition gradient and dilution experiments, and diamonds represent values reported by Aune and Timasheff, 1971, *Biochem.*, 10:1609-1617.

Although estimates of the association constants obtained by different treatments of the data are reasonably self-consistent, values of log K_{11} were systematically 5–10-fold lower at all pH values than those reported earlier by

Vincent and Lazdunski (1973, *Eur. J. Biochem.*, 38:365-372), although the pH dependence is approximately the same. There are at least two possible explanations for the discrepancy.

First, the difference may be due to the difference in buffers employed in the two studies. The equilibrium constants presented by Vincent and Lazdunski were calculated as the ratio of directly measured association and dissociation rate constants, and dissociation rates were measured in a buffer containing 50 mM $CaCl_2$ in addition to NaCl. Substitution of Ca^{2+} for Na^{+} has been found to significantly enhance noncovalent associations of a variety of other proteins, including the self-association of chymotrypsin (Aune et al., 1971, *Biochem.*, 10:1617-1622; Rivas et al., 1994, *Biochem.*, 33:2341-2348; Yu et al., 1997, *EMBO J.*, 16:5455-5463; Rivas et al., 1999, *Biochem.*, 38:9379-9388). We believe this is likely to be the case for associations between chymotrypsin and BPTI as well.

Second, discrepancy between the actual equilibrium constant and the apparent equilibrium constant calculated via the ratio of association and dissociation rate constants may arise if the complex AB exists as a mixture of two isomers in rapid equilibrium, only one of which undergoes dissociation. The actual and apparent equilibrium constants will then differ by a constant reflecting the fraction of AB existing as the dissociating isomer at equilibrium.

The global modeling of multiple experiments did not significantly increase the precision of results obtained from analysis of the combined data. We believe this is partly due to the fact that the individual one-component experiments and the two-component composition gradient experiments were carried out at different times, using different protein solutions. In addition to possible differences between the solutions used in the individual experiments, the possibility of slight changes in instrumental sensitivity also exists. The light scattering detector was recalibrated periodically—typically once every two months—but not before each experiment.

Nevertheless, the success of the composition gradient light scattering technique described herein in resolving association equilibria may be attributed to the high information content of the composition-dependent scattering profile, which becomes evident when the contributions of individual species to the total light scattering profile are quantified. See, for example, Fig. 14B. The dependence of equilibrium concentrations of individual macrosolute species—and hence the total scattering profile—upon overall solution composition as it varies along the f_A

coordinate is extremely sensitive to differences between alternative reaction schemes and small changes in equilibrium constants within a particular scheme.

Example 4

Detection and Characterization of Indefinite Self-Associations

In this Example, we demonstrate a system comprising a liquid dispensing instrument and light scattering and concentration detectors can rapidly detect and characterize the indefinite linear self-association of a single macromolecular species and precisely determine standard free energies of successive addition of monomer to growing oligomers.

Materials

FtsZ, a prokaryotic cytoskeleton protein, prepared in a buffer containing 50 mM Tris-HCl + 50 mM KCl + 0.1 mM GDP + 5 mM $MgCl_2$, pH 7.5 (Rivas et al., 2000, *J. Biol. Chem.*, 275:11740-11749), was a gift from Dr. Germán Rivas, CIB-CSIC. Protein concentrations were determined from the absorbance at 280 nm using the following standard values for absorbance in OD units/cm pathlength for a 1 g/l solution: FtsZ, 0.345. Refractive increments were determined as described in Example 1 and found to be equal to 0.185 ± 0.003 ml/g at 20°C. Immediately before light scattering measurement, solutions were prefiltered and centrifuged as described in Example 1. Measurements of light scattering were carried out at 20°C.

Instrumentation

Experiments were conducted utilizing the instrumentation described in Example 2. A schematic of the instrumentation is shown in Figure 2.

Experimental Procedure

Experiments on a single protein were performed as described in Example 1. Briefly, a solution of the protein was loaded into one reservoir (222) and buffer was loaded into the second reservoir (224). The composition gradient was obtained by incrementally increasing the flow rate from one syringe and simultaneously decreasing the flow rate from the second syringe to maintain a constant total flow rate. At user specified intervals, the syringe pumps were halted to ensure

equilibration of the mixture. In this fashion, a stepwise gradient of increasing or decreasing protein concentration was established.

Data Analysis

Raw data obtained from the experiments was processed as described in Example 1 and saved as files of $\{w_{tot}, <R>/K\}$, where w_{tot} denotes the total concentration of protein in units of g/L and $<R>/K$ denotes the Rayleigh ratio averaged from data obtained by multiple detectors scaled to an optical constant K as defined in Example 1. All calculations were performed automatically using scripts and functions written and executed in MATLAB (Mathworks, Natick, MA).

In a solution containing a single protein component, absorbance data were converted to values of w_{tot} as described in Example 1. To calculate the value of R/K as a function of w_{tot}, a model specifying the equilibrium concentrations of all macromolecular solute species present in detectable quantity as a function of the total w/v concentration of each protein component was constructed. The general procedure for constructing the model is described in Example 2.

In the presence of GDP and Mg, FtsZ has been shown to undergo indefinite self-association to form linear oligomers (Rivas et al., 2000, *J. Biol. Chem.*, 275:11740-11749). Accordingly, the following set of equilibrium constants for stepwise addition of monomer to oligomer were defined for $i \geq 2$,

$$K_i = \frac{c_i}{c_{i-1} c_1}, \tag{25}$$

where c_i denotes the molar concentration of i-mer, and $K_1 \equiv 1$. The condition of conservation of mass was expressed as

$$c_{tot} = w_{tot} / M_1 = \sum_i i c_i = \sum_i i Q_i c_1^i, \tag{26}$$

where M_1 is the molar mass of monomeric protein, and

$$Q_i = \prod_{j-1}^{i} K_i . \tag{27}$$

According to the theory of isoenthalpic linear indefinite self-association of identical subunits (Chatelier et al., 1987, *Biophys. Chem.*, 28:121-128), the value of K_i should gradually decrease with increasing oligomer size and approach a constant value, denoted K_∞, in the limit of large oligomer size. Therefore, we defined a unitless scaling factor $F_i \equiv K_i/K_\infty$ for $i > 1$, such that $F_2 > 1$ and $F_i \to 1$ as $i \to \infty$, and $F_1 \equiv 1$. Then

$$Q_i = K_\infty^{i-1} Z_i, \tag{28}$$

where

$$Z_i = \prod_{j-1}^{i} F_j .$$

We now define the unitless protein concentrations $c_{tot}^* = K_\infty c_{tot}$ and $c_1^* = K_\infty c_1$. It follows then from Eqs. 26-28 that

$$c_{tot}^* = \sum_i i Z_i c_1^{*i} . \tag{29}$$

Given an experimental value of w_{tot} and test values of M_1, K_∞, and the Z_i, Eq. 29 may be solved numerically for the value of c_1^*. Then from Eq. 28,

$$c_i = \frac{Z_i}{K_\infty} c_1^{*i}, \tag{30}$$

and the scaled Rayleigh ratio may be calculated according to

$$R/K = \sum M_i^2 c_i = \frac{M_1^2}{K_\infty} \sum i^2 Z_i c_1^{*i} . \tag{31}$$

It is evident that the above model, which contains an arbitrary number of independently variable F_i or the equivalent Z_i, must be simplified. This can be done by specifying a simple empirical functional dependence of F_i upon i obeying the

conditions set out in the above definition of F_i. It was shown previously that so long as the final calculated dependence of experimental signal upon total protein concentration is in agreement with experiment, the underlying distribution of species is independent of the precise form of the empirical function employed (Rivas et al., 2000, *J. Biol. Chem.*, 275:11740-11749). We therefore selected one of the empirical forms utilized by Rivas et al. for $i > 1$:

$$F_i = 1 + \frac{J-1}{(i-1)^{\alpha}}. \qquad (31)$$

Thus, specification of the test values of the empirical parameters J and α together with K_∞ and M_1 permits the calculation of R/K as a function of w_{tot}. In the numerical calculation, sums over species indicated above were evaluated up to $i = 100$. It was verified that all series converged well before this limit.

Results

Replicate dilution experiments were performed on the FtsZ protein solution as described in Example 1. The experimentally measured dependence of $<R>/K$ upon w_{tot} is plotted in Fig. 16. The model described by Eqs. 25-31 was fit to the experimental data, fixing the value of M_1 at 40,500 (Rivas et al., 2000. *J. Biol. Chem.*, 275:11740-11749). The open circles in Fig. 16 represent experimental data. The curve in Fig. 16 was calculated from an inverse-decay model using any of several sets of correlated parameter values leading to identical fits of the data. A broad variety of combinations of the parameters K_∞, J, and α were found capable of fitting the data to within experimental precision, indicating that these parameters are highly correlated, and hence may not be determined individually. One of the (identical) calculated best-fits is plotted in Fig. 16 together with the data.

Stepwise equilibrium constants and standard free energies of stepwise association were calculated as functions of oligomer size using each of the sets of model parameters that fits the data to within experimental precision. In Figs. 17A and 17B, calculated values are plotted for a number of such parameter sets. In Fig. 17A, K_i for addition of monomer to form i-mer is plotted as a clue of i. In Fig. 17B, G_i for addition of monomer to form i-mer is plotted as a clue of i. The open circles in Figs. 17A and 17B were calculated using the experimentally observed

dependence of $<R>/K$ on w_{tot}. It is evident that stepwise equilibrium constants and free energies of monomer association so calculated are uniquely determined to high precision, independent of the particular parameter set, so long as that parameter set is capable of fitting the experimental data to within experimental precision.

Also plotted in Figs. 17A and 17B are comparable values obtained from analysis of sedimentation equilibrium experiments conducted on FtsZ under nominally identical experimental conditions (Rivas et al., 2000. *J. Biol. Chem.*, 275:11740-11749). The dashed-line curve was calculated using parameters obtained by modeling sedimentation equilibrium data as described in Rivas et al. In view of the fact that the two sets of measurements were conducted more than five years apart by different investigators using entirely different techniques, the agreement between the stepwise free energies obtained by the two methods (generally better than ~0.1 RT) is particularly impressive. Figs. 17A and B further demonstrates that the systems and methods of the invention are capable of rapidly acquiring information about macromolecular associations that is comparable in scope and resolution to that currently obtainable from sedimentation equilibrium.

D. Software Code and Scripts

Tables 3-11 provided below include example software code and scripts that can be used to control the systems disclosed herein.

Table 3 - H541GR2.PAS

```
program H541GR2;
   {
    3/2/05 - v. 2
             original valve configuration
             add extra volume at end

    1/31/05 - new step dispensing algorithm for 541c
              new valve configuration
              250 ul syringes

    1/31/05 - modification for new valves - mirror image of left,
              right configuration

    1/31/05 - modifications of HAMGRD10 for Hamilton 541c

    12/7/04 - allow optional extra amount of 0%A or 100%A in
gradient

    3/11/04 - optional continuous gradient (no refill, but optional
           delay between steps)
```

```
    1/2/04 - slow down speed of refilling syringes

    12/09/03 - allow reverse autogradient, from 100 to 0

    12/09/03 - allow following choice:
       1. straight 0% injection
       2. straight 100% injection
       3. auto gradient - from 0 to 100 in 5% steps and
                  selectable aliquot sizes and time increments

    12/05/03 - modification to deliver gradient with gradual
increases of 5%,
      allow for pause between steps or no pause, allow for
variation in
      aliquot size

    9/24/03 - modification of hamimix8 to deliver linear gradient
of
      solute in syringe A from 0 to 100% over a time interval of
ca. 2 min

    5/22/03 - modif of hamimix5 and hamimix6 to allow for
programmed
      dispensing of samples for wyatt

    4/18/03 - modification of hamirean

    version of 04/15/03 - hamirean
    realtime control of 941

    uses async.tpu -- unit for control of serial port

    }

uses dos,crt,async;

type
   buffer = string[80];

var
   aliquotvol: double;
   i,errorcode,nL,nR,nreps,locationL,locationR,distanceL,
      distanceR, speed, step, nsteps, tpercentA,
      tcollecttime, exitflag, tvol, pauseflag,
        aliquottime, expel: integer;
   c,choice: char;
   nextadr,myadr,echo,reply,comstring,snL,snR,sdistL,sdistR,
        comstring1,comstring2,freestring, goodstring, yesstring,
        nostring, speedstr: buffer;
   collecttime: array[1..30] of integer;
   percentA, totvol: array[1..30] of double;

{*****************************************************************}

  procedure sendstring(outstring:buffer);

  var c: char;
      i: integer;
      echo: boolean;

  begin
```

```
  outstring := 'a' + outstring + #13;       {terminate with CR}

  for i := 1 to length(outstring) do
      begin
        c := outstring[i];
        Async_Send(c)
      end;

  repeat
        echo := Async_Buffer_Check(c);
  until echo = false;
  delay(50);

  end;

{*****************************************************************}

procedure recstring(var instring:buffer);
   {receives a string terminating in <CR> from the serial port}

   var scount: integer;
       a: char;

   begin
   scount := 0;
   instring := '';
   a := #0;
   while ((scount<5000) and (a<>#13)) do
      begin
      inc(scount);
      if async_buffer_check(a) then
         begin
         instring := instring + a;
         scount := 0;
         end;
      end;
   end;

{*****************************************************************}

function BUSY: boolean;
   {true if 541 is busy, false if otherwise}

   begin
   sendstring('T1');
   recstring(echo);
   recstring(reply);
   if reply=freestring then busy := false
                       else busy := true;
   end;

{*****************************************************************}

procedure COMMAND(outstr:buffer);
   {sends command string to 941 repeatedly until processed}

var reply, reply1, reply2: buffer;

  begin
```

```
    repeat until not busy;
       sendstring(outstr);
    repeat until not busy;

  end;

{*****************************************************************}

procedure QUERY(outstr:buffer; var reply: buffer);
   {sends command string to 941 repeatedly until processed}

var reply1, reply2: buffer;

  begin

       repeat until not busy;
       sendstring(outstr);
       recstring(reply1);
       recstring(reply2);
       reply := copy(reply2, 2, length(reply2) - 1);
             {strips ACK from reply}
  end;

{*****************************************************************}

procedure INIT_PUMP;
    {initialize 941}

begin
   async_init;
   if not async_open(1,9600,'O',7,1) then        {open COM port}
      begin
      writeln('**ERROR: Async_Open failed');
      halt;
      end;

   sendstring('1a');            {initialize 541}
   recstring(nextadr);
   recstring(myadr);

   command('BYSM1CYSM1R');      { set precision to 2000 steps }

 end;  {initpump}

{*****************************************************************}

procedure GETELTIME(var eltime: longint);

var hour, minute, sec, sec100: word;

begin

   gettime(hour, minute, sec, sec100);
   eltime := 3600*hour + 60*minute + sec;

end;

{*****************************************************************}

procedure GETLOC(i: integer; var loc: integer);
   { returns absolute location of syringe L (1) or R (2) }
```

```
var locstring: buffer;
    err: integer;

begin

   if i = 1 then
     query('bYQP', locstring)
   else
     query('cYQP', locstring);
   val(locstring, loc, err);

end;

{*****************************************************************}

procedure REFILL_SYRINGES;

{ fill both syringes }

begin

   command('BIM2000S10CIM2000S10R');
      repeat until not keypressed;

end;

{*****************************************************************}

procedure EMPTY_SYRINGES(valve:integer);

{ empty syringe to waste (1) or reservoir (2) }

begin

   if valve = 1 then
      command('BOM0S10COM0S10R')
   else
      command('BIM0S10CIM0S10R');

end;

{*****************************************************************}

procedure DISPENSESTEP(stepvol,percentA: double; steptime:
integer);
    { allowed values of stepvol = 100, 150, 200, 250
      allowed values of percentA = 0, 5, 10, 15, ... , 100 }

var eltimestart, eltimenow, deltime, eltimewas: longint;
    volperstep, totsteps, injecttime: double;
    Rsteps, Lsteps, Rspeed, Lspeed: integer;
    sRsteps, sLsteps, sRspeed, sLspeed: buffer;

begin

   geteltime(eltimestart);
   eltimewas := eltimestart;
   injecttime := stepvol/20.0;
   volperstep := 0.125;
   totsteps := stepvol/volperstep;
```

```
   Rsteps := round(0.01*percentA*totsteps);
   Lsteps := round(totsteps - Rsteps);
   if Rsteps = 0 then
      Rspeed := 100
   else
      Rspeed := round(injecttime*2000/Rsteps);
   if Lsteps = 0 then
      Lspeed := 100
   else
      Lspeed := round(injecttime*2000/Lsteps);
   str(Rsteps,sRsteps);
   str(Lsteps,sLsteps);
   str(Rspeed,sRspeed);
   str(Lspeed,sLspeed);
   comstring := 'BOD' + sLsteps + 'S' + sLspeed + 'COD' +
        sRsteps + 'S' + sRspeed + 'R';
   command(comstring);

   refill_syringes;

      repeat
      geteltime(eltimenow);
      deltime := eltimenow - eltimestart;
      if eltimenow > eltimewas then
         begin
            write(^M);
            clreol;
            write(deltime);
         eltimewas := eltimenow;
         end;
   until deltime > steptime;
   write(^M);

end;

{*****************************************************************}

procedure ADDVOL(percentA,volume: integer);

var syringeful, leftover, leftovertime: integer;

begin
      syringeful := volume div 250;
   leftover := volume mod 250;
   leftovertime := (leftover*15) div 250;
   if syringeful > 0 then
        for i := 1 to syringeful do
           dispensestep(250,percentA,15);
   if leftover > 0 then
           dispensestep(leftover,percentA,leftovertime);
end;

{*****************************************************************}

procedure AUTOGRADIENT(direction: integer);
  {  direction = 1 is gradient up, direction <> 1 is gradient down
}

var addextra1, addextra2: boolean;
    extravol1, extravol2: integer;
```

```
begin

   write('Enter intermediate step volume (ul) [100, 150, 200,
250]',
       ^J^M,'   and minimum step time (sec): ');
   readln(aliquotvol,aliquottime);
   write('Pause between steps? (Y/N): ');
   readln(choice);
   if upcase(choice) = 'Y' then
      pauseflag := 1
   else
      pauseflag := 0;

   write('Allow extra injection at start of gradient? (Y/N): ');
   readln(choice);
   addextra1 := false;
   if upcase(choice) = 'Y' then
      begin
      addextra1 := true;
      write('    Enter extra volume before first fraction (ul): ');
      readln(extravol1);
            end;

   write('Allow extra injection at end of gradient? (Y/N): ');
   readln(choice);
   addextra2 := false;
   if upcase(choice) = 'Y' then
      begin
      addextra2 := true;
      write('    Enter extra volume after last fraction (ul): ');
      readln(extravol2);
            end;

   nsteps := 21;

   for i := 1 to 21 do
      begin
      totvol[i] := aliquotvol;
      collecttime[i] := aliquottime;
      if direction = 1 then
            percentA[i] := 5*(i-1)
      else
            percentA[i] := 5*(21-i);
      end;

   clrscr;

   if addextra1 = true then
      if direction = 1 then
         addvol(0,extravol1)
      else
         addvol(100,extravol1);

   for step := 1 to nsteps do
      begin
      writeln('Step ', step, ': dispensing ', totvol[step]:3:0, '
of ',
               percentA[step]:3:0, '%A');
      dispensestep(totvol[step],percentA[step],collecttime[step]);
      refill_syringes;
      if pauseflag = 1 then
```

```
        begin
        write('Hit enter to continue ...');
          readln;
          end;
       end;

   if addextra2 = true then
      if direction = 1 then
         addvol(100,extravol2)
      else
         addvol(0,extravol2);

   repeat until not busy;
   writeln('*** end of step gradient ***');

end;

{*****************************************************************}

procedure CONTINGRADIENT(direction: integer);

     { direction = 1  0 -> 100% A
                 = 2  100 -> 0% A

     { total volume injected in gradient is 1000 steps each syringe
}

   const
         astep: array[1..21] of integer =

(0,5,10,14,19,24,29,33,38,43,48,52,57,62,67,71,76,81,86,90,95);
      bstep: array[1..21] of integer =

(95,90,86,81,76,71,67,62,57,52,48,43,38,33,29,24,19,14,10,5,0);

   var deltime, istep, nL, nR: integer;
   var eltimewas, eltimenow: longint;

begin

   locationL := 1000;
   locationR := 1000;
      write('Delay time between steps (sec): ');
   readln(deltime);

   for istep := 1 to 21 do
      begin
      if direction = 1 then
         begin
         nL := astep[istep];
         nR := bstep[istep];
         end
      else
         begin
         nL := bstep[istep];
         nR := astep[istep];
         end;
      str(nL,snL);
      str(nR,snR);
      comstring := 'BOD' + snL + 'S' + speedstr + 'CID' + snR +
                          'S' + speedstr + 'R';
```

```
      repeat until not busy;
            sendstring(comstring);
      locationL := locationL - nL;
      locationR := locationR - nR;
            repeat until not busy;
      geteltime(eltimewas);
            repeat
         geteltime(eltimenow);
      until (eltimenow - eltimewas) > deltime;
   end;

   repeat until not busy;
   writeln('*** end of continuous gradient ***');
   writeln('Hit enter to refill syringes');
   readln;
   refill_syringes;

end;

{*****************************************************************}

begin {main}

   {get things initialized}

   freestring := #6 + #127 + #13;          { <ACK> + ASCII 127 +
<CR> }

   clrscr;

   init_pump;

   write('Hit enter to initialize pump');
   readln;

   command('XOR');
{
   write('Enter 1 to empty syringes to waste (check valves!)',^J^M,
            '      2 to empty syringes to reservoir: ');
   repeat
      readln(expel);
   until expel in [1,2];

   empty_syringes(expel);
}
   write('Hit enter to refill syringes: ');
   readln;

   refill_syringes;

   write('Reset valves for experiment, then hit enter ...');
   readln;

   speedstr := '60';

   repeat

      repeat
         repeat until not busy;
            write(^J^M,'Select (1) to inject 0%', ^J^M,
                              '       (2) to inject 100%,', ^J^M,
```

```
                    '       (3) step gradient 0 -> 100,',^J^M,
                                    '       (4) step gradient 100
-> 0,',^J^M,
                    '       (5) cont. gradient 0 -> 100,',^J^M,
                    '       (6) cont. gradient 100 -> 0',^J^M,
                                    '       (7) to exit program:
');
            readln(choice);
      until choice in ['1','2','3','4','5','6','7'];

      case choice of
             '1': dispensestep(250,0,10);
                 '2': dispensestep(250,100,10);
                 '3': autogradient(1);      { refill gradient up }
             '4': autogradient(2);      { refill gradient down }
             '5': contingradient(1);    { continuous gradient up}
             '6': contingradient(2);    { continuous gradient down}
      end;

   until choice = '7';

   write('Hit enter to empty syringes back into reservoirs.');
   readln;

   empty_syringes(2);

end.
```

Table 4 - process_0503ab.m

```
% process_0503ab - new data filter algorithm 05/03/05
%    see lines 517 ff
% process_1124ab - display scattering, absorbance in select data
window
% process_1110ab - same as 1105, but no sigma in saved data set
% process_1105ab - calculates uncertainty of processed data
% process_1028ab - corrects defaults, plots, selects region from
absorbance plot
% process_1001ab - put concentration determination ahead of region
selection
% process_0917ab - for processing of 100%A to 100%B gradients
% process_0729t - option of thinning data
% process_0729 - uncertainty of regression coefficients
% process_0728 - scale graphics of selected concentration, 90o
scattering
% process_0504 - add choice of sensitivities
% process_0416
%   - add preview Mw vs w, Mw vs log w plots after data truncation
%   - add dataset number to give unique filenames for saving
different regions of each wyatt data set
% process_0401: option to average data from all detectors at each
time point
% process_0305: introduce new default inputs
% process_0227 - save processed data sorted on concentration
2/27/04
%    note: new value for Ainst (Attri)
% process_0213 - save,load normalization table
% process_0204 - do continuous processing of data over selected
```

```
range rather than peak averages
%   can elect to skip alignment if raw alignment is satisfactory

% version 0204 - iterative fit, eliminating outliers

%    1. zoom into region of interest

% process_1211 - use new 2dpolyfit
% process_1205 - align scattering, absorbance plots
% process_1204 - modification to load data from unedited Wyatt
export file
% process_1203 - display absorbance plot
% process_1201 - 12/01/03
% modification of process_0520 to input AUX 1 (absorbance) and
automatically
% convert to concentration

% process_0520
% modification of process_0519 to average normalization over all
dilutions
%          instead of selecting a single dilution
% process_0519
% revision of process_0221 for microbatch measurements

% process_0221
% Processing of Wyatt batch data
%     1. noise filtering
%  2. baseline subtraction
%  3. normalization
%  4. removal of bad points by inspection
%  5. fitting of good data by 2D polynomial

% process_0220 - saves .uda file containing only yanked points
% process_0212 - Weida corrections for instrumental constants
% process_0211y - normalize using one dilution
% process_0211 - load concentration series from file
%     1. use lowest concentration of protein in dextran to calculate
normalization
%           coefficients
%     2. use total refractive index of solution to calculate
refraction,reflection for instrument
%           optical constant
% process_0205 - load normalization coefficients, concentrations
from file
% process_0131.m
% new normalization constants derived from analysis of new Damien
data 01/31/03
% new data load algorithm

warning off MATLAB:divideByZero

% theta for scintillation vials
% theta = [22.5 28 32 38 44 50 57 64 72 81 90 99 108 117 126 134 141
147];

% theta for flow cell - aqueous solvent
% note: first 2 detectors inaccessible in flow cell
theta = [0 0 14.5 25.9 34.8 42.8 51.5 60 69.3 79.7 90 100.3 110.7
121.2 132.2 142.5 152.5 163.3];
xang = (sin(pi*theta/360)).^2;       % angle function in Zimm-Debye
equation
```

```
nwater = 1.330;      % H2O
dndwdex = 0.145;       % ml/g
% dndwprot = 0.19;    % ml/g  05/30 - assign later

% refractive index of scintillation vial glass
% nglass = 1.505;       % Wyatt value (Miles Weida)

% refractive index of K5 flow cell glass
 nglass = 1.519;    % temporary value until we get value from Weida
 fresnel2 = 1 - ((nglass - 1)/(nglass + 1))^2;

% calibration constant of Damien data (from toluene - lowest
sensitivity?)
% Ainst = 9.296e-4;       % 01/31/03

% calibration constant of new Minton data
% Ainst = 1.068e-5;      % obtained from methanol measurement
05/15/03 - highest sensitivity

% calibration constant of new Attri data
% Ainst = 9.3e-6;        % obtained from methanol 11/28/03 - high
sensitivity

% calibration constant of Attri 03/02/04
% Ainst = 1.04e-5;

% calibration constant of new Attri data - for each of three
sensitivity levels
% Ainsttable = [1.07e-3 5.0e-5 1.064e-5];        % obtained from
methanol 11/28/03 - high sensitivity

% calibration constants measured 2 Aug 2004
Ainsttable = [1.016e-3 4.72e-5 9.74e-6];        % uncertainties:
[0.02e-6 0.05e-5 (unmeasurable)]

% calibration constants measured March 1 2005
% Ainsttable = [1.17e-3 5.39e-5 1.12e-5];

choice = menu('Select instrument sensitivity','Low
(1X)','Intermediate (21x)','High (100X)');
Ainst = Ainsttable(choice);

homefolder = pwd;
workpath = uigetfolder('Select data folder');
cd(workpath);

dataset = 0;      % an identifier for processed data sets

while 1

   close all;
   clear fulldata rawdata bcrawdata normcoef conc clear dndwprot
absorbancedat


   [filename,pathname] = uigetfile('*.txt','Select Wyatt export
file');
   [path,fileprefix,ext,ver] = fileparts(filename);

   if filename==0
```

```
      close all;
      cd(homefolder);
      return
   end

   dataset = dataset + 1;

   fullfilename = [pathname filename];
   disp(['*** Data set ' num2str(dataset) ': ' fullfilename]);

   % new 12/04/03

   in = fopen(fullfilename,'rt');           % open file
   for j=1:9
       line = fgets(in);                    %skip first 9 lines
(text headers)
   end
   formatstr = '%g %g %g %g %g %g %g %g ';
   formatstr = [formatstr formatstr formatstr];
   fulldata = fscanf(in,formatstr,[24 inf]);   % read 24 numbers per
line until eof
   fulldata = fulldata';                       % transpose to
reconstruct data
   fclose(in);

   [path,name,ext,ver] = fileparts(fullfilename);

   % create diary file

   diaryname = filenamegen;
   diaryfile = [diaryname '.prc'];      % prc is 'process' file
   diary(diaryfile);

   disp('Processing of Wyatt gradient file - process_1028ab');
   disp(['Diary file: ' diaryfile]);
   disp(['Working folder: ' workpath]);
   disp(['Data file name: ' filename]);
   disp(' ');

   % ************ New 08/27/04 - option for thinning data
*************************

   disp([num2str(length(fulldata)) ' data points']);
   thinchoice = menu('Thin data?','Yes','No');

   if thinchoice == 1
       thinfact = input('Enter factor by which data to be thinned
(integer): ');
       thinfact = round(thinfact);
       fulldata = fulldata(1:thinfact:end,:);
       savechoice = menu('Save thinned data?','Yes','No');
       if savechoice == 1
           savename = [name '_thin' num2str(thinfact)];
           fullsavename = [path '\' savename ext];
           save(fullsavename,'fulldata','-ascii');
           disp(['Data thinned by factor of ' num2str(thinfact)
'saved as ' fullsavename]);
       end
   end

   % ********* following modifications are for solutions without
```

```
dextran (07/29/04) ***********

   % eliminate menu choice of protein or dextran as main solute

   % solutechoice = menu('Solute is:','protein','dextran');

   % if solutechoice==1
      disp('Solute is protein');
      dndwprot = input('Enter dn/dw of protein (null default value =
0.185): ');
      if isempty(dndwprot)
          dndwprot = 0.185;
      end

      % eliminate selection of dextran concentration

   %   wdex = input('Enter dextran concentration (g/l) (default =
0): ');
   %   if isempty(wdex)
   %       wdex = 0;
   %   end
   % else
   %   disp('Solute is dextran');
   % end

   wdex = 0;

   % ************ end elimination of dextran modifications 07/29/04
*************************

   % new 12/01 - channel 19 is AUX1 (absorbance)

   rawdata = [];
   for i = 1:19
      rawdata(:,i) = fulldata(:,i+3);
   end

   fig1 = figure('Units', 'normal', 'Position', [.15 .05 .7 .4],
'PaperPositionMode','auto');
   rawplot = plot(rawdata(:,11),'.','MarkerSize',1);
   hold on;
   border = axis;
   border(3) = border(3) - 0.01;
   newymax = input('Enter new ymax, or null to leave unchanged: ');
   if ~isempty(newymax)
      border(4) = newymax;
   end
   axis(border);

   % select baseline, calculate excess scattering

   title('Click on left end of baseline, or click right button to
assume zero baseline');
   [x,y,button] = ginput(1);
   if button>1
      base(1:19) = 0;
      bcrawdata = rawdata;                                % bcrawdata is
baseline-corrected raw data
   else
      leftb1 = round(x);
      markerx = [leftb1 leftb1];
```

```
      markery = [border(3) border(4)];
      plot(markerx,markery,'g');
      title('Click on right end of baseline');
      [x,y] = ginput(1);
      rightb1 = round(x);
      markerx = [rightb1 rightb1];
      markery = [0 border(4)];
      plot(markerx,markery,'g');
        labelx = 0.5*leftb1 + 0.5*rightb1;
      labely = 0.05*markery(2);
      text(labelx,labely,'B','Color','g');
      base = dustfilter19(rawdata(leftb1:rightb1,:));
      for i = 1:19
         bcrawdata(:,i) = rawdata(:,i) - base(i);
      end
   end

   % ------------------------------------------------------------------
-------

   % new - determine concentrations of 100%A and 100%B - Sept 17
2004

   fig1b = figure('Units', 'normal', 'Position', [.1 .05 .8 .4],
'PaperPositionMode','auto');
   absorbance = bcrawdata(:,19);
   plot(absorbance,'b');    % plot absorbance
   ylabel('Absorbance');
   hold on
   border = axis;
   ymark = [border(3) border(4)];
   title('Mark region corresponding to
100%A','FontSize',14,'Color','r');
   [t(1),y] = ginput(1);
   xmark = [t(1) t(1)];
   plot(xmark,ymark,'r');
   [t(2),y] = ginput(1);
   xmark = [t(2) t(2)];
   plot(xmark,ymark,'r');
   Aplotrange = round(t(1)):round(t(2));
   AbsorbA = mean(absorbance(Aplotrange));

   % calculate stdev of scattering of pure A for each detector - 5
nov 04

   for i = 1:18
       sigscatA(i) = 1.414*std(bcrawdata(Aplotrange,i));
   end

   % multiplication by sqrt(2) takes into account that bcrawdata is
the difference between two
   % measurements

   plot([t(1) t(2)],[AbsorbA AbsorbA],'r');
   text(t(1),1.05*AbsorbA,'100%A','Color','r');

   % allow for uncertainty of concentration determination - 5 nov 04

   % QA = minput('Enter absorbance of 1 g/l A, sigma (%): ',2);
   QA = minput('Enter absorbance of 1 g/l A: ',1);    % new 24 nov
   epsA = QA(1);
```

```
    % sigepsA = epsA*0.01*QA(2);
    sigepsA = 0;
    wAo = AbsorbA/epsA;
    sigwAo = wAo*sigepsA/epsA;
    disp(['   wAo = ' num2str(wAo) '   sigma(wAo) = '
num2str(sigwAo)]);
    title('Mark region corresponding to
100%B','FontSize',14,'Color','r');
    [t(3),y] = ginput(1);
    xmark = [t(3) t(3)];
    plot(xmark,ymark,'r');
    [t(4),y] = ginput(1);
    xmark = [t(4) t(4)];
    plot(xmark,ymark,'r');
    Bplotrange = round(t(3)):round(t(4));
    AbsorbB = mean(absorbance(Bplotrange));

    % calculate stdev of scattering of pure B for each detector - 5
nov 04

    for i = 1:18
        sigscatB(i) = 1.414*std(bcrawdata(Bplotrange,i));
    end

    % multiplication by sqrt(2) takes into account that bcrawdata is
the difference between two
    % measurements

    plot([t(3) t(4)],[AbsorbB AbsorbB],'r');
    text(t(3),1.05*AbsorbB,'100%B','Color','r');

    % allow for uncertainty of concentration determination - 5 nov 04

    % QB = minput('Enter absorbance of 1 g/l B, sigma (%): ',2);
    QB = minput('Enter absorbance of 1 g/l B: ',1);    % new 24 nov
    epsB = QB(1);
    % sigepsB = epsB*0.01*QB(2);
    sigepsB = 0;
    wBo = AbsorbB/epsB;
    sigwBo = wBo*sigepsB/epsB;
    disp(['   wBo = ' num2str(wBo) '   sigma(wBo) = '
num2str(sigwBo)]);

    close(fig1b);

    % ------------------------------------------------------------------
-----------------

    % new 11/24/04 - select region to analyze from combined
absorbance, 90o scattering plot

    fig1c = figure('Units', 'normal', 'Position', [.1 .05 .8 .4],
'PaperPositionMode','auto');

    newxplotmin = min([round(t(2)) round(t(3))]);
    newxplotmax = max([round(t(2)) round(t(3))]);
    newxplotrange = newxplotmin:newxplotmax;
    newyplotmin = -0.05;
    newyplotmax = 1.05;
    minabsorb = min(bcrawdata(newxplotrange,19));
    maxabsorb = max(bcrawdata(newxplotrange,19));
```

```
    scaledabsorbdat = (bcrawdata(newxplotrange,19) -
minabsorb)/(maxabsorb - minabsorb);
    minscat = min(bcrawdata(newxplotrange,11));
    maxscat = max(bcrawdata(newxplotrange,11));
    scaledscatdat = (bcrawdata(newxplotrange,11) - minscat)/(maxscat
- minscat);
    plot(newxplotrange,scaledabsorbdat,'k');
    hold on
    plot(newxplotrange,scaledscatdat,'b');
    axis([newxplotmin newxplotmax newyplotmin newyplotmax]);

    ymark = [newyplotmin newyplotmax];
    title('Mark start of region to be analyzed');
    [tplot(1),abmark(1)] = ginput(1);
    xmark = [tplot(1) tplot(1)];
    plot(xmark,ymark,'m');
    title('Mark end of region to be analyzed');
    [tplot(2),abmark(2)] = ginput(1);
    xmark = [tplot(2) tplot(2)];
    plot(xmark,ymark,'m');
    xtext = (9*tplot(1) + tplot(2))/10;
    ytext = (ymark(1) + ymark(2))/2;
    text(xtext,ytext,'data range','Color','m');
    input ('Print if desired, then hit enter to continue ...');
    close(fig1b);

    tplot(1) = round(tplot(1));
    tplot(2) = round(tplot(2));
    tplotrange = tplot(1):tplot(2);
    npoints = length(tplotrange);
    newborder = [tplot(1) tplot(2) border(3) border(4)];
    axis(newborder);

    % ---------------------- new 07/27/04 ---------------------------
-------------
    % scale plots of conc, 90o scattering

    concdat = bcrawdata(:,19)/(1.05*max(bcrawdata(:,19)));    % scaled
absorbance data
    scatdat = bcrawdata(:,11);
    maxconcplot = max(concdat(tplotrange));
    minconcplot = min(concdat(tplotrange));
    minscatplot = min(scatdat(tplotrange));
    maxscatplot = max(scatdat(tplotrange));

    close all;
    fig2 = figure('Units', 'normal', 'Position', [.1 .05 .8 .4],
'PaperPositionMode','auto');

 while 1      % graphic matching loop

    normscatdat = minconcplot + (maxconcplot - minconcplot)* ...
                          (scatdat - minscatplot)/(maxscatplot -
minscatplot);
    scatplot = plot(tplotrange,normscatdat(tplotrange),'k');
% plot normalized 90o scattering
    hold on
    concplot = plot(tplotrange,concdat(tplotrange),'r');
    legend('90^o scatter','absorbance',0);

    choice = menu('Is additional scaling necessary?','Yes','No');
```

```
    if choice==1
        title('Mark position of scattering signal to match to conc
signal');
        [xmatch,y] = ginput(1);
        xmatch = round(xmatch);
        maxscatplot = scatdat(xmatch);
        delete(scatplot);
    else
        break
    end

  end     % scaling loop

   choice = menu('Is additional alignment necessary?','Yes','No');
   if choice==1
        title('Mark first time benchmark in scatter
data','FontSize',14,'Color','r');
        [tscat(1),y] = ginput(1);
        title('Mark first time benchmark in concentration
data','FontSize',14,'Color','r');
        [tconc(1),y] = ginput(1);
        disp(['First point offset: ' num2str(tconc(1) - tscat(1))]);
        title('Mark second time benchmark in scatter
data','FontSize',14,'Color','r');
        [tscat(2),y] = ginput(1);
        title('Mark second time benchmark in concentration
data','FontSize',14,'Color','r');
        [tconc(2),y] = ginput(1);
        disp(['Second point offset: ' num2str(tconc(2) -
tscat(2))]);
        offset = round(mean(tconc - tscat));
        disp(['Average offset: ' num2str(offset)]);
    else
        offset = 0;
    end

   clear selectshiftconcdat selectabsorbance selectnormscatdat
selectbcrawdata nsolvent sigscat fA
   clear fresnel Acscc Kopt normdata RoK normcoef detweight select
signormdata sigRoK signormcoef

   for ipt = 1:npoints
       selectshiftconcdat(ipt) = concdat(tplot(1) - 1 + ipt +
offset);
       selectabsorbance(ipt) = bcrawdata((tplot(1) - 1 + ipt +
offset),19);
       selectnormscatdat(ipt) = normscatdat((tplot(1) - 1 + ipt));
   end

   close(fig2);
   fig3 = figure('Units', 'normal', 'Position', [.1 .05 .8 .4],
'PaperPositionMode','auto');

   selectbcrawdata = bcrawdata(tplotrange,:);
   normscatplot = plot(selectnormscatdat,'k');
   hold on
   shiftconcplot = plot(selectshiftconcdat,'b');
   border = axis;
   axis([border(1) border(2) border(3)
1.05*max(selectshiftconcdat)]);
   legend('90^o scatter','aligned absorbance',0);
```

```
    not_ok = (selectbcrawdata >= 10);
    okcheck = sum(sum(not_ok));
    disp([ num2str(okcheck) ' data points exceed 10V']);
    input('Hit enter to continue ...');
    close(fig3);

    % calculate sigma of selectbcrawdata - 05 nov 04

    sigscat = [];
    fA = ((selectabsorbance - AbsorbB)/(AbsorbA - AbsorbB))';
    for iang = 1:18
        sigscat(:,iang) = fA*sigscatA(iang) + (1 -
fA)*sigscatB(iang);
    end

    % normalization

    conc = selectabsorbance/1000;    % units of g/ml
    nsolvent = nwater + dndwdex*wdex*0.001 + dndwprot*conc;
    fresnel = fresnel2*(1 - ((nglass - nsolvent)/(nglass +
nsolvent))^2)^2;
    Acscc = Ainst*nglass*nsolvent./fresnel;
    Kopt = 2.9e-6*dndwprot^2*nsolvent.^2;

            % calculate stddev of normcoef for each detector - 5 nov
04

   choice = menu('Normalization options','Self-normalize','Load
normalization data');
   for iang = 1:18
        normarray = selectbcrawdata(:,11)./selectbcrawdata(:,iang);
        normcoef(iang) = mean(normarray);
        if ((normcoef(iang) == Inf) | (normcoef(iang) == 0))
            normcoef(iang) = 0;
            detweight(iang) = 0;
        else
            signormcoef(iang) = std(normarray)/(sqrt(npoints));
            % stdev of mean
            if signormcoef(iang) > .03*normcoef(iang)
                detweight(iang) = 0;
            else
                detweight(iang) = 1;
            end
        end
    end
    if choice==1
        choice2 = menu('Save self-normalization data?','Yes','No');
        if choice2==1
            normfilename = [fileprefix '.nrm'];
            savedata = [normcoef' signormcoef'];
            save(normfilename,'savedata','-ascii');
        end
    else
      [normfilename,pathname] = uigetfile('*.nrm','Select
normalization file');
      normcoefdat = load(normfilename);
      normcoef = normcoefdat(:,1);
      signormcoef = normcoefdat(:,2);
    end
```

```
    normdata = [];
    signormdata = [];
    RoK = [];
    sigRoK = [];

    for iang = 1:18
         normdatadet = normcoef(iang)*selectbcrawdata(:,iang);
         varnormdatadet =
normdatadet.^2.*((signormcoef(iang)/normcoef(iang))^2 + ...
                 (sigscat(:,iang)./selectbcrawdata(:,iang)).^2);
         signormdatadet = sqrt(varnormdatadet);
         normdata = [normdata normdatadet];
         signormdata = [signormdata signormdatadet];
         RoK = [RoK (Acscc.*normdatadet'./Kopt)'] ;
         sigRoK = [sigRoK (Acscc.*signormdatadet'./Kopt)'];
         disp(['detector: ' num2str(iang) '   normcoef = '
num2str(normcoef(iang)) ...
               '   detweight = ' num2str(detweight(iang))]);
    end

% new 05/03/05 - eliminate polynomial smoothing
% for each time point do a linear lsq fit of xang vs RoK for
detectors 8:16
% calculate best fit straight line, residuals for all detectors 4:18
% set weight = 0 for a detector if the squared residual > 3 * mean
square residual for the fit

w = zeros(size(RoK));   % initialize weight array

xdat = [];
ydat = [];
zdat = [];

for ipt = 1:npoints
    Pt = polyfit(xang(8:16),RoK(ipt,8:16),1);
    yc = polyval(Pt,xang(4:18));
    ssrt = (yc - RoK(ipt,4:18)).^2;
    mssrt = mean(ssrt(5:13));   % only over those residuals in the
fit data set
    wt = (ssrt > 3*mssrt).*zeros(size(ssrt)) + (ssrt <=
3*mssrt).*ones(size(ssrt));
    w(ipt,4:18) = wt;
    for iang = 4:18
        if detweight(iang) == 1
            if w(ipt,iang) > 0
                xdat = [xdat; fA(ipt)];
                ydat = [ydat; xang(iang)];
                zdat = [zdat; RoK(ipt,iang)];
            end
        end
    end
end

% ------------------------------------------------

disp(['Number of data points: ' num2str(length(xdat))]);

%plot filtered data

   xmin = min(xdat);
   xmax = max(xdat);
```

```
    zmin = min(zdat);
    zmax = max(zdat);
    xrange = xmax - xmin;
    zrange = zmax - zmin;
    xplotmin = xmin - 0.05*xrange;
    xplotmax = xmax + 0.05*xrange;
    yplotmin = 0;
    yplotmax = 1;
    zplotmin = zmin - 0.05*zrange;
    zplotmax = zmax + 0.05*zrange;

    close   % previous figure

    figure('Position',[362    43   380
273],'PaperPositionMode','auto');
    plot3(xdat, ydat, zdat, 'ko','MarkerFaceColor','c');
    boundaries = [xplotmin xplotmax yplotmin yplotmax zplotmin
zplotmax];
    grid off;
    axis(boundaries);
    grid on;

    xlabel('fA');
    ylabel('sin^2(\theta/2)');
    zlabel('R/K');
    view(48,18);

    rotate3d on;

    title(fixtitle(filename));
    input('Hit enter to proceed ...');

       plotfile = myinputdlg({'Save plot file name
[none]'},'',1,{''},[350 200]);
       if ~strcmp(plotfile,'')
          hgsave (gcf,plotfile);
          disp(['Figure saved as ' plotfile '.fig']);
       else
          disp(' ');
       end

    response = questdlg('Save filtered data set?
[No]','','Yes','No','No');
    if strcmp(response,'Yes')
        clear outdat soutdat
        response = questdlg('Save angular dependence?
[Condense]','','Yes','No','Condense','Condense');
        if strcmp(response,'Yes')
            outdat = [xdat ydat zdat];
            fullfilename = [fileprefix '_' datestr(now,'mmmddyy')
'_' num2str(dataset) '_3d.uda'];
        elseif strcmp(response,'Condense')
            indat = sortrows([xdat zdat],1);
            outdat = avgang2(indat);
            fullfilename = [fileprefix '_' datestr(now,'mmmddyy')
'_' num2str(dataset) '.uda'];
        else
            % outdat = [xdat zdat wdat];
            outdat = [xdat zdat];
            fullfilename = [fileprefix '_' datestr(now,'mmmddyy')
'_' num2str(dataset) '.uda'];
```

```
        end
        soutdat = sortrows(outdat,1);
        save (fullfilename , 'soutdat', '-ascii');
        disp(['Filtered processed data set saved in: '
fullfilename]);
    end

% clean up, get out

      close;
      diary off;
      disp([diaryfile ' closed.']);

response = questdlg('Another data set? [No]','', ...
        'Yes','No','No');
if strcmp(response,'No')
   close all;
   break
end

end    % of data set loop
```

Table 5 - process_0729f.m

```
% process_0729ff  (substitute getdir for getfolder)
% process_0729f
% process_0729t - option of thinning data
% process_0729 - uncertainty of regression coefficients
% process_0728 - scale graphics of selected concentration, 90o
scattering
% process_0504 - add choice of sensitivities
% process_0416
%   - add preview Mw vs w, Mw vs log w plots after data truncation
%   - add dataset number to give unique filenames for saving
different regions of each wyatt data set
% process_0401: option to average data from all detectors at each
time point
% process_0305: introduce new default inputs
% process_0227 - save processed data sorted on concentration
2/27/04
%    note: new value for Ainst (Attri)
% process_0213 - save,load normalization table
% process_0204 - do continuous processing of data over selected
range rather than peak averages
%   can elect to skip alignment if raw alignment is satisfactory

% version 0204 - iterative fit, eliminating outliers

%    1. zoom into region of interest

% process_1211 - use new 2dpolyfit
% process_1205 - align scattering, absorbance plots
% process_1204 - modification to load data from unedited Wyatt
export file
% process_1203 - display absorbance plot
% process_1201 - 12/01/03
% modification of process_0520 to input AUX 1 (absorbance) and
automatically
% convert to concentration
```

```
% process_0520
% modification of process_0519 to average normalization over all
dilutions
%          instead of selecting a single dilution
% process_0519
% revision of process_0221 for microbatch measurements

% process_0221
% Processing of Wyatt batch data
%     1. noise filtering
%  2. baseline subtraction
%  3. normalization
%  4. removal of bad points by inspection
%  5. fitting of good data by 2D polynomial

% process_0220 - saves .uda file containing only yanked points
% process_0212 - Weida corrections for instrumental constants
% process_0211y - normalize using one dilution
% process_0211 - load concentration series from file
%     1. use lowest concentration of protein in dextran to calculate
normalization
%           coefficients
%     2. use total refractive index of solution to calculate
refraction,reflection for instrument
%           optical constant
% process_0205 - load normalization coefficients, concentrations
from file
% process_0131.m
% new normalization constants derived from analysis of new Damien
data 01/31/03
% new data load algorithm

warning off MATLAB:divideByZero

% theta for scintillation vials
% theta = [22.5 28 32 38 44 50 57 64 72 81 90 99 108 117 126 134 141
147];

% theta for flow cell - aqueous solvent
% note: first 2 detectors inaccessible in flow cell
theta = [0 0 14.5 25.9 34.8 42.8 51.5 60 69.3 79.7 90 100.3 110.7
121.2 132.2 142.5 152.5 163.3];
xang = (sin(pi*theta/360)).^2;       % angle function in Zimm-Debye
equation

nwater = 1.330;      % H2O
dndwdex = 0.145;       % ml/g
% dndwprot = 0.19;    % ml/g  05/30 - assign later

% refractive index of scintillation vial glass
% nglass = 1.505;       % Wyatt value (Miles Weida)

% refractive index of K5 flow cell glass
 nglass = 1.519;    % temporary value until we get value from Weida
 fresnel2 = 1 - ((nglass - 1)/(nglass + 1))^2;

% calibration constant of Damien data (from toluene - lowest
sensitivity?)
% Ainst = 9.296e-4;       % 01/31/03

% calibration constant of new Minton data
```

```
% Ainst = 1.068e-5;      % obtained from methanol measurement
05/15/03 - highest sensitivity

% calibration constant of new Attri data
% Ainst = 9.3e-6;        % obtained from methanol 11/28/03 - high
sensitivity

% calibration constant of Attri 03/02/04
% Ainst = 1.04e-5;

% calibration constant of new Attri data - for each of three
sensitivity levels
% Ainsttable = [1.07e-3 5.0e-5 1.064e-5];        % obtained from
methanol 11/28/03 - high sensitivity

% calibration constants measured 2 Aug 2004
Ainsttable = [1.016e-3 4.72e-5 9.74e-6];        % uncertainties:
[0.02e-6 0.05e-5 (unmeasurable)]

% calibration constants measured March 1 2005
% Ainsttable = [1.17e-3 5.39e-5 1.12e-5];

choice = menu('Select instrument sensitivity','Low
(1X)','Intermediate (21x)','High (100X)');
Ainst = Ainsttable(choice);

homefolder = pwd;
workpath = uigetdir('c:\matlab\wyatt','Select data folder');
cd(workpath);

dataset = 0;      % an identifier for processed data sets

while 1

   close all;
   clear fulldata rawdata bcrawdata normcoef conc clear dndwprot
absorbancedat


   [filename,pathname] = uigetfile('*.txt','Select Wyatt export
file');
   [path,fileprefix,ext,ver] = fileparts(filename);

   if filename==0
      close all;
      cd(homefolder);
      return
   end

   dataset = dataset + 1;

   fullfilename = [pathname filename];
   disp(['*** Data set ' num2str(dataset) ': ' fullfilename]);

   % new 12/04/03

   in = fopen(fullfilename,'rt');              % open file
   for j=1:9
       line = fgets(in);                       %skip first 9 lines
(text headers)
   end
```

```
    formatstr = '%g %g %g %g %g %g %g %g ';
    formatstr = [formatstr formatstr formatstr];
    fulldata = fscanf(in,formatstr,[24 inf]);    % read 24 numbers per
line until eof
    fulldata = fulldata';                        % transpose to
reconstruct data
    fclose(in);

    [path,name,ext,ver] = fileparts(fullfilename);

    % create diary file

    diaryname = filenamegen;
    diaryfile = [diaryname '.prc'];       % prc is 'process' file
    diary(diaryfile);

    disp('Processing of Wyatt gradient file - process_0729t');
    disp(['Diary file: ' diaryfile]);
    disp(['Working folder: ' workpath]);
    disp(['Data file name: ' filename]);
    disp(' ');

    % ************* New 08/27/04 - option for thinning data
*************************

    disp([num2str(length(fulldata)) ' data points']);
    thinchoice = menu('Thin data?','Yes','No');

    if thinchoice == 1
        thinfact = input('Enter factor by which data to be thinned
(integer): ');
        thinfact = round(thinfact);
        fulldata = fulldata(1:thinfact:end,:);
        savechoice = menu('Save thinned data?','Yes','No');
        if savechoice == 1
            savename = [name '_thin' num2str(thinfact)];
            fullsavename = [path '\' savename ext];
            save(fullsavename,'fulldata','-ascii');
            disp(['Data thinned by factor of ' num2str(thinfact)
'saved as ' fullsavename]);
        end
    end

    % ********* following modifications are for solutions without
dextran (07/29/04) ***********

    % eliminate menu choice of protein or dextran as main solute

    % solutechoice = menu('Solute is:','protein','dextran');

    % if solutechoice==1
       disp('Solute is protein');
       dndwprot = input('Enter dn/dw of protein (null default value =
0.185): ');
       if isempty(dndwprot)
           dndwprot = 0.185;
       end

       % eliminate selection of dextran concentration

    %    wdex = input('Enter dextran concentration (g/l) (default =
```

```
0): ');
   %   if isempty(wdex)
   %       wdex = 0;
   %   end
   % else
   %   disp('Solute is dextran');
   % end

   wdex = 0;

   % ************ end elimination of dextran modifications 07/29/04
*************************

   % new 12/01 - channel 19 is AUX1 (absorbance)

   rawdata = [];
   for i = 1:19
      rawdata(:,i) = fulldata(:,i+3);
   end

   fig1 = figure('Units', 'normal', 'Position', [.15 .05 .7 .4],
'PaperPositionMode','auto');
   rawplot = plot(rawdata(:,11),'.','MarkerSize',1);
   hold on;
   border = axis;
   border(3) = border(3) - 0.01;
   newymax = input('Enter new ymax, or null to leave unchanged: ');
   if ~isempty(newymax)
      border(4) = newymax;
   end
   axis(border);

   % select baseline, calculate excess scattering

   title('Click on left end of baseline, or click right button to
assume zero baseline');
   [x,y,button] = ginput(1);
   if button>1
      base(1:19) = 0;
      bcrawdata = rawdata;                                    % bcrawdata is
baseline-corrected raw data
   else
      leftb1 = round(x);
      markerx = [leftb1 leftb1];
      markery = [border(3) border(4)];
      plot(markerx,markery,'g');
      title('Click on right end of baseline');
      [x,y] = ginput(1);
      rightb1 = round(x);
      markerx = [rightb1 rightb1];
      markery = [0 border(4)];
      plot(markerx,markery,'g');
        labelx = 0.5*leftb1 + 0.5*rightb1;
      labely = 0.05*markery(2);
      text(labelx,labely,'B','Color','g');
      base = dustfilter19(rawdata(leftb1:rightb1,:));
      for i = 1:19
         bcrawdata(:,i) = rawdata(:,i) - base(i);
      end
   end
```

```
   % new 1/20 - select region of interest

   border = axis;
   ymark = [border(3) border(4)];
   title('Mark start of region to be analyzed');
   [tplot(1),y] = ginput(1);
   xmark = [tplot(1) tplot(1)];
   plot(xmark,ymark,'r');
   title('Mark end of region to be analyzed');
   [tplot(2),y] = ginput(1);
   tplot(1) = round(tplot(1));
   tplot(2) = round(tplot(2));
   tplotrange = tplot(1):tplot(2);
   npoints = length(tplotrange);
   newborder = [tplot(1) tplot(2) border(3) border(4)];
   axis(newborder);

   % ---------------------- new 07/27/04 ---------------------------
-------------
   % scale plots of conc, 90o scattering

   concdat = bcrawdata(:,19)/(1.05*max(bcrawdata(:,19)));   % scaled
absorbance data
   scatdat = bcrawdata(:,11);
   maxconcplot = max(concdat(tplotrange));
   minconcplot = min(concdat(tplotrange));
   minscatplot = min(scatdat(tplotrange));
   maxscatplot = max(scatdat(tplotrange));

   close(fig1);
   fig2 = figure('Units', 'normal', 'Position', [.1 .05 .8 .4],
'PaperPositionMode','auto');

 while 1     % graphic matching loop

   normscatdat = minconcplot + (maxconcplot - minconcplot)* ...
                        (scatdat - minscatplot)/(maxscatplot -
minscatplot);
   scatplot = plot(tplotrange,normscatdat(tplotrange),'k');
% plot normalized 90o scattering
   hold on
   concplot = plot(tplotrange,concdat(tplotrange),'r');
   legend('90^o scatter','absorbance',0);

   choice = menu('Is additional scaling necessary?','Yes','No');
   if choice==1
        title('Mark position of scattering signal to match to conc
signal');
        [xmatch,y] = ginput(1);
        xmatch = round(xmatch);
        maxscatplot = scatdat(xmatch);
        delete(scatplot);
    else
        break
    end

  end     % scaling loop

   choice = menu('Is additional alignment necessary?','Yes','No');
   if choice==1
        title('Mark first time benchmark in scatter
```

```
data','FontSize',14,'Color','r');
        [tscat(1),y] = ginput(1);
        title('Mark first time benchmark in concentration
data','FontSize',14,'Color','r');
        [tconc(1),y] = ginput(1);
        disp(['First point offset: ' num2str(tconc(1) - tscat(1))]);
        title('Mark second time benchmark in scatter
data','FontSize',14,'Color','r');
        [tscat(2),y] = ginput(1);
        title('Mark second time benchmark in concentration
data','FontSize',14,'Color','r');
        [tconc(2),y] = ginput(1);
    -   disp(['Second point offset: ' num2str(tconc(2) -
tscat(2))]);
        offset = round(mean(tconc - tscat));
        disp(['Average offset: ' num2str(offset)]);
    else
        offset = 0;
    end

   clear selectshiftconcdat selectabsorbance selectnormscatdat
selectbcrawdata nsolvent
   clear fresnel Acscc Kopt normdata RoK normcoef detweight

   for ipt = 1:npoints
       selectshiftconcdat(ipt) = concdat(tplot(1) - 1 + ipt +
offset);
       selectabsorbance(ipt) = bcrawdata((tplot(1) - 1 + ipt +
offset),19);
       selectnormscatdat(ipt) = normscatdat((tplot(1) - 1 + ipt));
   end

   close(fig2);
   fig3 = figure('Units', 'normal', 'Position', [.1 .05 .8 .4],
'PaperPositionMode','auto');

   selectbcrawdata = bcrawdata(tplotrange,:);
   normscatplot = plot(selectnormscatdat,'k');
   hold on
   shiftconcplot = plot(selectshiftconcdat,'b');
   border = axis;
   axis([border(1) border(2) border(3)
1.05*max(selectshiftconcdat)]);
   legend('90^o scatter','aligned absorbance',0);

   % normalization

   not_ok = (selectbcrawdata >= 10);
   okcheck = sum(sum(not_ok));
   disp([ num2str(okcheck) ' data points exceed 10V']);
   input('Hit enter to continue ...');

   excoef = input('Enter A280 of 1 g/l solution (default = 1.0): ');
   if isempty(excoef)
       excoef = 1.0;
   end
   conccoef = 1/excoef;
   conc = conccoef*selectabsorbance/1000;    % units of g/ml
   nsolvent = nwater + dndwdex*wdex*0.001 + dndwprot*conc;
   fresnel = fresnel2*(1 - ((nglass - nsolvent)/(nglass +
nsolvent))^2)^2;
```

```
    Acscc = Ainst*nglass*nsolvent./fresnel;
    Kopt = 2.9e-6*dndwprot^2*nsolvent.^2;

    choice = menu('Normalization options','Self-normalize','Load
normalization data');
    for iang = 1:18
         normarray = selectbcrawdata(:,11)./selectbcrawdata(:,iang);
         normcoef(iang) = mean(normarray);
         if ((normcoef(iang) == Inf) | (normcoef(iang) == 0))
             normcoef(iang) = 0;
             detweight(iang) = 0;
         else
             stddevnormcoef = std(normarray)/(sqrt(npoints));
             if stddevnormcoef > .03*normcoef(iang)
                 detweight(iang) = 0;
             else
                 detweight(iang) = 1;
             end
         end
     end
     if choice==1
         choice2 = menu('Save self-normalization data?','Yes','No');
         if choice2==1
             normfilename = [fileprefix '.nrm'];
             savedata = normcoef';
             save(normfilename,'savedata','-ascii');
         end
     else
       [normfilename,pathname] = uigetfile('*.nrm','Select
normalization file');
       normcoef = load(normfilename);
     end

    normdata = [];
    RoK = [];

    for iang = 1:18
         normdatadet = normcoef(iang)*selectbcrawdata(:,iang);
         normdata = [normdata normdatadet];
         RoK = [RoK (Acscc.*normdatadet'./Kopt)'] ;
         disp(['detector: ' num2str(iang) '   normcoef = '
num2str(normcoef(iang)) ...
                '   detweight = ' num2str(detweight(iang))]);
    end

    outdat = [];
    for iang = 4:18
         if detweight(iang) == 1
            for ipt = 1:npoints
                 outdat = [outdat; conc(ipt) xang(iang)
RoK(ipt,iang)];
            end
         end
    end


    response = questdlg('Save data set? [No]','','Yes','No','No');
     if strcmp(response,'Yes')
          clear soutdat

          soutdat = sortrows(outdat,1);
```

```
        fullfilename = [fileprefix num2str(dataset) '.uda'];
        save (fullfilename , 'soutdat', '-ascii');
        disp(['Processed data set saved in: ' fullfilename]);
    end

% fitting module modified from lsfit.m
% linear least squares solution to 2D polynomial
% x = concentration
% y = sin^2(theta/2)
% z = RoK

% ---------------------------------------- New fit Dec 11 03 --------
------------------

clear xdat ydat zdat wdat

response = 'Yes';

while strcmp(response,'Yes')      % parameter set loop (same model,
same data set)

xdat = outdat(:,1);
ydat = outdat(:,2);
zdat = outdat(:,3);
wdat = ones(size(xdat));

disp(['Number of data points: ' num2str(length(xdat))]);

while 1   % outlier removal iterative fit loop

   %plot data

   xmin = 0;
   xmax = max(xdat);
   ymin = 0;
   ymax = max(ydat);
   zmin = 0;
   zmax = max(zdat);
   xrange = xmax - xmin;
   yrange = ymax - ymin;
   zrange = zmax - zmin;
   xplotmin = 0;
   xplotmax = xmax + 0.05*xrange;
   yplotmin = 0;
   yplotmax = ymax + 0.05*yrange;
   zplotmin = 0;
   zplotmax = zmax + 0.05*zrange;

   figure('Position',[50 50 600 230],'PaperPositionMode','auto');
   subplot(1,2,1);
   plot3(xdat, ydat, zdat, 'ko','MarkerFaceColor','c');
      hold on;
   boundaries = [xplotmin xplotmax yplotmin yplotmax zplotmin
zplotmax];
   grid off;
   axis(boundaries);
   grid on;

   xlabel('conc (g/ml)');
   ylabel('sin^2(\theta/2)');
   zlabel('R/K');
```

```
   view(48,18);

   rotate3d on;

   xplotmin = boundaries(1);
   xplotmax = boundaries(2);
   yplotmin = boundaries(3);
   yplotmax = boundaries(4);
   zplotmin = boundaries(5);
   zplotmax = boundaries(6);

   xplot = linspace(xplotmin, xplotmax, 10);
   yplot = linspace(yplotmin, yplotmax, 10);
   [xplotg,yplotg] = meshgrid(xplot,yplot);

   title(fixtitle(filename));
   hold off;

    clear Cx Dx;

    choice = input('Constrain baseline to zero? (Y/N) (default = Y):
','s');
    if (strcmp(upper(choice),'Y') | length(choice)==0)
        jmin = 1;
    else
        jmin = 0;
    end

      maxval = minput('Enter max order for conc dep, max order for
ang dep (default = 1,0): ',2);
    if sum(maxval==[0 0]) == 2
        jmax = 1;
        kmax = 0;
    else
          jmax = maxval(1);
          kmax = maxval(2);
    end

      [Cx, Dx] =
xtwodpolyfitd(xdat,ydat,zdat,wdat,jmin,jmax,0,kmax);          % column
vector
      Cx = Cx';     % row vector

      nbfpars = length(Cx);
      nfitpts = length(xdat);

      zcalc = xtwodpolyfunc(jmin,jmax,0,kmax,Cx,xdat,ydat);
      res = zdat - zcalc;
      sqrres = res.^2;
      DOF = nfitpts - nbfpars;
      bfsigsqr = sum(sqrres)/DOF;

      sigCx = sqrt(bfsigsqr*Dx);          % Appendix B8 of DG
Kleinbaum et al, Applied
                                          % Regression Analysis and
Other Multivariable
                                          % Methods
      Cxdisp = [];
      sigCxdisp = [];
      nrows = jmax - jmin + 1;                    % 1st row -
extrapolation to zero conc
```

```
      ncols = kmax + 1;              % 1st col - extrapolation to
zero angle
      for irow = 1:nrows
         Cxdisp = [Cxdisp; Cx((irow-1)*ncols + (1:ncols))];
         sigCxdisp = [sigCxdisp; sigCx((irow-1)*ncols + (1:ncols))];
            end

      disp(' ');
      disp('BEST FIT PARAMETERS');
      disp(Cxdisp);

      disp(' ');
      disp('SIGMA OF BEST FIT PARAMETERS');
      disp('(Calculated assuming random and normally distributed
residuals)');
      disp(sigCxdisp);

      meansqrres = mean(sqrres);
      newwdat = wdat.*(sqrres < 3*meansqrres);

      wssr = sum(wdat.*sqrres);
      wsstr = sprintf('\nBest fit WSSR: %6.4e',wssr);
      disp(wsstr);
      disp(['DOF: ' int2str(DOF)]);

      % plot best fit

        zplotg = xtwodpolyfunc(jmin,jmax,0,kmax,Cx,xplotg,yplotg);
      hold on;
      % if size(zplot,1)~=size(xplot,1)
      %    zplot = zplot';
      % end
      mesh(xplotg,yplotg,zplotg);
      black = zeros(64,3);
      colormap(black);
      hidden off;
      boundaries(6) = max([max(max(zplotg)) boundaries(6)]);
                  % increases height of plot if necessary
      axis(boundaries);
      [az,el] = view;

      % plot residuals of best fit

      resmin = min(res);
      resmax = max(res);
      resplotrange = resmax - resmin;
      resplotmin = resmin - 0.05*resplotrange;
      resplotmax = resmax + 0.05*resplotrange;
      zerosurf = zeros(size(xplotg));
      subplot(1,2,2);
      plot3(xdat,ydat,res,'ko','MarkerFaceColor','c');
      hold on;
      mesh(xplotg,yplotg,zerosurf);
      axis ([xplotmin xplotmax yplotmin yplotmax resplotmin
resplotmax]);
      hidden off;
      grid on;

      xlabel('conc');
      ylabel('sin^2(\theta/2)');
      zlabel('resid');
```

```
      view(az,el);

      rotate3d on;

      titlestring = ['Diary file: ' diaryfile];
            title(fixtitle(titlestring));

      choice = upper(input('Refit to data set without outliers?
','s'));
      if ~strcmp(choice,'Y')
          break
      end

      [newxdat,newydat,newzdat,newwdat] =
truncdata3(xdat,ydat,zdat,newwdat);

      disp(['Number of points: ' num2str(length(newxdat))]);

      clear xdat ydat zdat wdat

      xdat = newxdat;
      ydat = newydat;
      zdat = newzdat;
      wdat = newwdat;

  end   % outlier removal loop

      response = questdlg('Save calculated best fit?
[No]','','Yes','No','No');
      if strcmp(response,'Yes')
         outdata = [xdat ydat zdat zcalc res];
         [filename,pathname]=uiputfile('.cal','Best fit results
filename');
         fullfilename=[pathname filename];
         save fullfilename outdata -ascii;
         disp(['Calculated best fit saved in: ' fullfilename]);
      end

      plotfile = myinputdlg({'Save plot file name
[none]'},'',1,{''},[350 200]);
      if ~strcmp(plotfile{1},'')
         hgsave (gcf,plotfile{1});
         disp(['Figure saved as ' plotfile{1} '.fig']);
      else
         disp(' ');
      end

    response = questdlg('Save truncated data set?
[No]','','Yes','No','No');
    if strcmp(response,'Yes')
        clear outdat soutdat
        response = questdlg('Save angular dependence?
[Condense]','','Yes','No','Condense','Condense');
        if strcmp(response,'Yes')
            outdat = [xdat ydat zdat];
            fullfilename = [fileprefix '_' datestr(now,'mmmddyy')
'_' num2str(dataset) '.t3d'];
        elseif strcmp(response,'Condense')
            indat = sortrows([xdat zdat],1);
            outdat = avgang2(indat);
            fullfilename = [fileprefix '_' datestr(now,'mmmddyy')
```

```
'_' num2str(dataset) '.uda'];
            response2 = questdlg('View plots of Mw,app vs w, Mw,app
vs log w? [No]','','Yes','No','No');
            if strcmp(response2,'Yes')
                clear w Mwapp
                w = outdat(:,1);
                Mwapp = outdat(:,2)./outdat(:,1);
                mwplot = figure('Position',[187    50   677
234],'PaperPositionMode','auto');
                subplot(1,2,1)
                plot(w,Mwapp,'.');
                xlabel('w (g/ml)');
                ylabel('M_{W,app}');
                subplot(1,2,2)
                plot(log10(w),Mwapp,'.');
                xlabel('log w(g/ml)');
                ylabel('M_{W,app}');
                suptitle(fixtitle(fullfilename));
                disp('Edit, print, or save plots and then hit enter
to continue ...');
                pause;
                close(mwplot);
            end
        else
            outdat = [xdat zdat];
            fullfilename = [fileprefix '_' datestr(now,'mmmddyy')
'_' num2str(dataset) '.uda'];
        end
        soutdat = sortrows(outdat,1);
        save (fullfilename , 'soutdat', '-ascii');
        disp(['Truncated processed data set saved in: '
fullfilename]);
    end

   response = questdlg('Another fit to same data set? [No]','', ...
      'Yes','No','No');

end

% ------------------------------------- End of new fit Dec 11 ---
-----------------

% clean up, get out

      close;
      diary off;
      disp([diaryfile ' closed.']);

response = questdlg('Another data set? [No]','', ...
        'Yes','No','No');
if strcmp(response,'No')
   close all;
   break
end

end    % of data set loop
```

Table 6 - xf_AA_AB_scat_fA2.m

```
function y = xf_AA_AB_scat_fA(mode,P,x)

% xf_AA_AB_scat_fA 17 Feb 05, modification of
% xf_AplusB_scat_fA to allow dimerization of A
% fix calculation of wAtot, wBtot
% x = fA
% y = <R>/K
% P(1) - wAo (conc of A in fA = 1) ** a constant
% P(2) - wBo (conc of A in fB = 1) ** a constant
% P(3) = MA
% P(4) = MB
% P(5) = log Kab (M^-1)
% P(6) = log Kaa (M^-1)

global npars parname

if mode==0
   %---------------- parameter names and setup calcs ---------------
------

   npars = 6;
   parname = {'wAo (constant)','wBo (constant)','MA','MB','log Kab
(M^-1)' ...
       'log Kaa (M^-1)'};
   y = zeros(size(x));

else
   % -------------------- function calcs --------------------------
------

    fA = x;
    wAo = P(1);
    wBo = P(2);
    MA = P(3);
    MB = P(4);
    MAB = MA + MB;
    MAA = 2*MA;
    Kab = 10^P(5);          % inv molar units
    Kaa = 10^P(6);
    wAtot = wAo*fA;
    wBtot = wBo*(1 - fA);
    cAtot = wAtot/MA;
    cBtot = wBtot/MB;
    coef(1) = 2*Kaa*Kab;
    coef(2) = Kab + 2*Kaa;
    for i = 1:length(x)
        coef(3) = 1 + Kab*cBtot(i) - Kab*cAtot(i);
        coef(4) = -cAtot(i);
        cA = goodroot(coef,0,cAtot(i));
        cB = cBtot(i)/(1 + Kab*cA);
        cAA = Kaa*cA^2;
        cAB = Kab*cA*cB;
        RoK(i) = (MA^2*cA + MB^2*cB + MAB^2*cAB + MAA^2*cAA)/1e3;
            % g/ml units
    end
    y = RoK';

end
```

Table 7 - xf_AA_AB_scat_fA.m

```
function y = xf_AA_AB_scat_fA(mode,P,x)

% xf_AA_AB_scat_fA 17 Feb 05, modification of
% xf_AplusB_scat_fA to allow dimerization of A
% fix calculation of wAtot, wBtot
% x = fA
% y = <R>/K
% P(1) - wAo (conc of A in fA = 1) ** a constant
% P(2) - wBo (conc of A in fB = 1) ** a constant
% P(3) = MA
% P(4) = MB
% P(5) = log Kab (M^-1)
% P(6) = log Kaa (M^-1)

global npars parname

if mode==0
   %---------------- parameter names and setup calcs ---------------
------

   npars = 6;
   parname = {'wAo (constant)','wBo (constant)','MA','MB','log Kab
(M^-1)' ...
       'log Kaa (M^-1)'};
   y = zeros(size(x));

else
   % -------------------- function calcs ---------------------------
------

    fA = x;
    wAo = P(1);
    wBo = P(2);
    MA = P(3);
    MB = P(4);
    MAB = MA + MB;
    MAA = 2*MA;
    Kab = 10^P(5);          % inv molar units
    Kaa = 10^P(6);
    wAtot = wAo*fA;
    wBtot = wBo*(1 - fA);
    cAtot = wAtot/MA;
    cBtot = wBtot/MB;
    coef(1) = 2*Kaa*Kab;
    coef(2) = Kab + 2*Kaa;
    for i = 1:length(x)
        coef(3) = 1 + Kab*cBtot(i) - Kab*cAtot(i);
        coef(4) = -cAtot(i);
        cA = goodroot(coef,0,cAtot(i));
        cB = cBtot(i)/(1 + Kab*cA);
        cAA = Kaa*cA^2;
        cAB = Kab*cA*cB;
        RoK(i) = (MA^2*cA + MB^2*cB + MAB^2*cAB + MAA^2*cAA)/1e3;
            % g/ml units
    end
    y = RoK';

end
```

Table 8 - xf_AplusB_scat_fA.m

```
function y = xf_AplusB_scat_fA(mode,P,x)

% xf_AplusB_scat_fA
% fix calculation of wAtot, wBtot
% x = fA
% y = <R>/K
% P(1) - wAo (conc of A in fA = 1) ** a constant
% P(2) - wBo (conc of A in fB = 1) ** a constant
% P(3) = MA
% P(4) = MB
% P(5) = log Kab (M^-1)

global npars parname

if mode==0
   %---------------- parameter names and setup calcs ---------------
------

   npars = 5;
   parname = {'wAo (constant)','wBo (constant)','MA','MB','log Kab
(M^-1)'};
   y = zeros(size(x));

else
   % -------------------- function calcs --------------------------
------

    fA = x;
    wAo = P(1);
    wBo = P(2);
    MA = P(3);
    MB = P(4);
    MAB = MA + MB;
    Km = 10^P(5);           % inv molar units
    K = Km*MAB/(MA*MB);     % inv w/v units
    K1 = K*MA/MAB;
    K2 = K*MB/MAB;
    wAtot = wAo*fA;
    wBtot = wBo*(1 - fA);
    b = 1 + K1*wBtot - K2*wAtot;
    wA = (-b + sqrt(b.^2 + 4*K2*wAtot))/(2*K2);
    wB = wBtot./(1 + K2*wA);
    wAB = K*wA.*wB;
    RoK = (MA*wA + MB*wB + MAB*wAB)/1e3;  % (g/l -> g/ml)
    y = RoK;

end
```

Table 9 - xf_isodesmscat.m

```
function y = xf_isodesmscat(mode,P,x)

% xf_isodesmscat
% isodesmic scattering model
% x = w_tot (g/ml)
% y = R/K
% P(1) = M1
% P(2) = log K (M^-1)  stepwise K
```

```
global npars parname dataset ndatasets

if mode==0
   %---------------- parameter names and setup calcs ---------------
------

   npars = 2;
   parname = {'M1','log K2 (M^-1)'};
   y = zeros(size(x));

else
   % -------------------- function calcs --------------------------
------

    wtot = 1000*x;      % convert concentrations to g/l
    M1 = P(1);
    K = 10^P(2);    % M^-1
    ctot = wtot/M1;
    cstar = K*ctot;
    b = 2*cstar + 1;
    q = (b - sqrt(b.^2 - 4*cstar.^2))./(2*cstar);
    RoK = M1^2/K*(q + q.^2)./((1 - q).^3);
    y = RoK/1e3;     % convert to g/ml conc scale

end
```

Table 10 - xf_kinvdecay_scat.m

```
function y = xf_kinvdecay_scat(mode,P,x)

% xf_kinvdecay_scat
% quasi-isodesmic scattering model
% x = w_tot (g/ml)
% y = R/K
% P(1) = M1
% P(2) = log Kinf (M^-1)
% P(3) = J = K2/Kinf
% P(4) = alpha

global npars parname

if mode==0
   %---------------- parameter names and setup calcs ---------------
------

   npars = 4;
   parname = {'M1','log Kinf','J','alpha'};
   y = zeros(size(x));

else
   % -------------------- function calcs --------------------------
------

    wtot = 1000*x;      % convert concentrations to g/l
    M1 = P(1);
    Kinf= 10^P(2);    % M^-1
    J = P(3);
    alpha = P(4);
    cstartot = Kinf*wtot/M1;
```

```
    global Z imax ctottemp

    imax = 100;
    Z(1) = 1;
    for i = 2:imax
        KoKinf(i) = 1 + (J - 1)/((i-1)^alpha);
        Z(i) = Z(i-1)*KoKinf(i);
    end

    j = 1:imax;
    jsq = j.^2;

    for i=1:length(x)
        ctottemp = cstartot(i);
        c1star = fzero(@decayfunc,[0 ctottemp]);
        RoK(i) = M1^2*sum(jsq.*Z.*c1star.^j)/Kinf;
    end

    y = RoK'/1e3;    % convert to g/ml

end

% ------------------------------------------------------------------

function y = decayfunc(x)

% mass conservation equation for decay models

global imax Z ctottemp

ssum = x;
for i=2:imax
   ssum = ssum + i*Z(i)*x^i;
end
y = ssum - ctottemp;
```

Table 11 - xf_mondimtrim_scat.m

```
function y = xf_mondimtrim_scat(mode,P,x)

% xf_kinvdecay_scat
% quasi-isodesmic scattering model
% x = w_tot (g/ml)
% y = R/K
% P(1) = M1
% P(2) = log K2 (M^-1)
% P(3) = beta = K3/K2

global npars parname

if mode==0
   %---------------- parameter names and setup calcs ---------------
------

   npars = 3;
   parname = {'M1','log K2','beta'};
   y = zeros(size(x));

else
```

```
   % -------------------- function calcs --------------------------
------

    wtot = 1000*x;      % convert concentrations to g/l
    M1 = P(1);
    K2= 10^P(2);    % M^-1
    beta = P(3);
    cstartot = K2*wtot/M1;

    global Z imax ctottemp

    imax = 3;
    j = 1:imax;
    jsq = j.^2;
    Z = [1 1 beta];

    for i=1:length(x)
       ctottemp = cstartot(i);
       c1star = fzero(@mascon,[0 ctottemp]);
       RoK(i) = M1^2*sum(jsq.*Z.*c1star.^j)/K2;
    end

    y = RoK'/1e3;    % convert to g/ml

end

% ---------------------------------------------------------------------

function y = mascon(x)

% mass conservation equation for decay models

global imax Z ctottemp

ssum = x;
for i=2:imax
   ssum = ssum + i*Z(i)*x^i;
end
y = ssum - ctottemp;
```

The various embodiments described above are provided by way of illustration only and should not be construed to limit the invention. Those skilled in the art will readily recognize various modifications and changes that may be made to the present invention without following the example embodiments and applications illustrated and described herein, and without departing from the true spirit and scope of the present invention, which is set forth in the following claims.

It should be noted that, as used in this specification and the appended claims, the singular forms "a", "an" and "the" include plural referents unless the content clearly dictates otherwise. It should also be noted that the term "or" is generally employed in its sense including "and/or" unless the content clearly dictates otherwise.

All publications and patent applications in this specification are indicative of the level of ordinary skill in the art to which this disclosure pertains. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated by reference.

What is claimed is:

1. A system for detecting macromolecular interactions in solution, the system comprising:

a dispenser module configured to dispense at least one solution comprising at least one macromolecule;

a detector configured to simultaneously measure light scattering and concentration associated with the macromolecule in the solution; and

wherein the dispenser module is configured to vary a concentration of the macromolecule in the solution over time as the solution is delivered to the detector.

2. The system of claim 1, wherein the detector includes a first detector configured to measure light scattering, and a second detector configured to measure concentration, and wherein the first and second detectors are positioned in parallel, so that the first and second detectors take simultaneous measurements of the light scattering and concentration of the macromolecule in solution.

3. The system of claim 2, wherein the dispenser module comprises first and second syringes to dispense first and second solutions.

4. The system of claim 3, wherein a rate of dispensing of the first syringe is changed with respect to a rate of dispensing of the second syringe to vary a concentration of the first solution with respect to the second solution over time.

5. The system of claim 3, wherein the first solution comprises the macromolecule, and the second solution comprises a solvent.

6. The system of claim 3, wherein the first solution comprises a first macromolecule, and the second solution comprises a second macromolecule.

7. The system of claim 3, further comprising a mixer module coupled to the first and second syringes, wherein the mixer module is configured to mix the first and second solutions.

8. The system of claim 2, further comprising a splitter module configured to split the solution into parallel streams having a balanced flow rate, a first of the parallel streams being delivered to the first detector, and a second of the parallel streams being delivered to the second detector.

9. The system of claim 2, further comprising a computer system in communication with one or more of the first and second detectors, the computer system being configured to analyze data collected from the first and second detectors.

10. The system of claim 9, wherein the computer system is configured to generate a model of association for the macromolecule based on the data from the first and second detectors.

11. A method for detecting macromolecular interactions in solution, the method comprising:

providing a detector;

dispensing at least one solution comprising at least one macromolecule;

varying a concentration of said at least one macromolecule in the solution over time as the solution is delivered to the detector; and

measuring simultaneously light scattering and concentration of the solution using the detector.

12. The method of claim 11, wherein the step of providing the detector further comprises providing first and second detectors.

13. The method of claim 12, further comprising:

positioning the first and second detectors in parallel;

splitting the solution into parallel streams with a balanced flow rate;

delivering a first of the parallel streams to a first detector to measure light scattering associated with the macromolecule in the solution; and

delivering a second of the parallel streams to a second detector to measure concentration associated with the macromolecule in the solution.

14. The method of claim 11, wherein the step of dispensing further comprises using first and second syringes to dispense first and second solutions.

15. The method of claim 14, further comprising mixing the first and second solutions from the first and second syringes.

16. The method of claim 15, further comprising changing a rate of dispensing of the first syringe with respect to a rate of dispensing of the second syringe to vary a concentration of the first solution with respect to the second solution over time.

17. The method of claim 11, wherein a first solution comprises a macromolecule and a second solution comprises a solvent.

18. The method of claim 11, wherein a first solution comprises a first macromolecule and a second solution comprises a second macromolecule.

19. The method of claim 11, further comprising generating a model of association for the macromolecule based on the measurements of the first detector and the second detector.

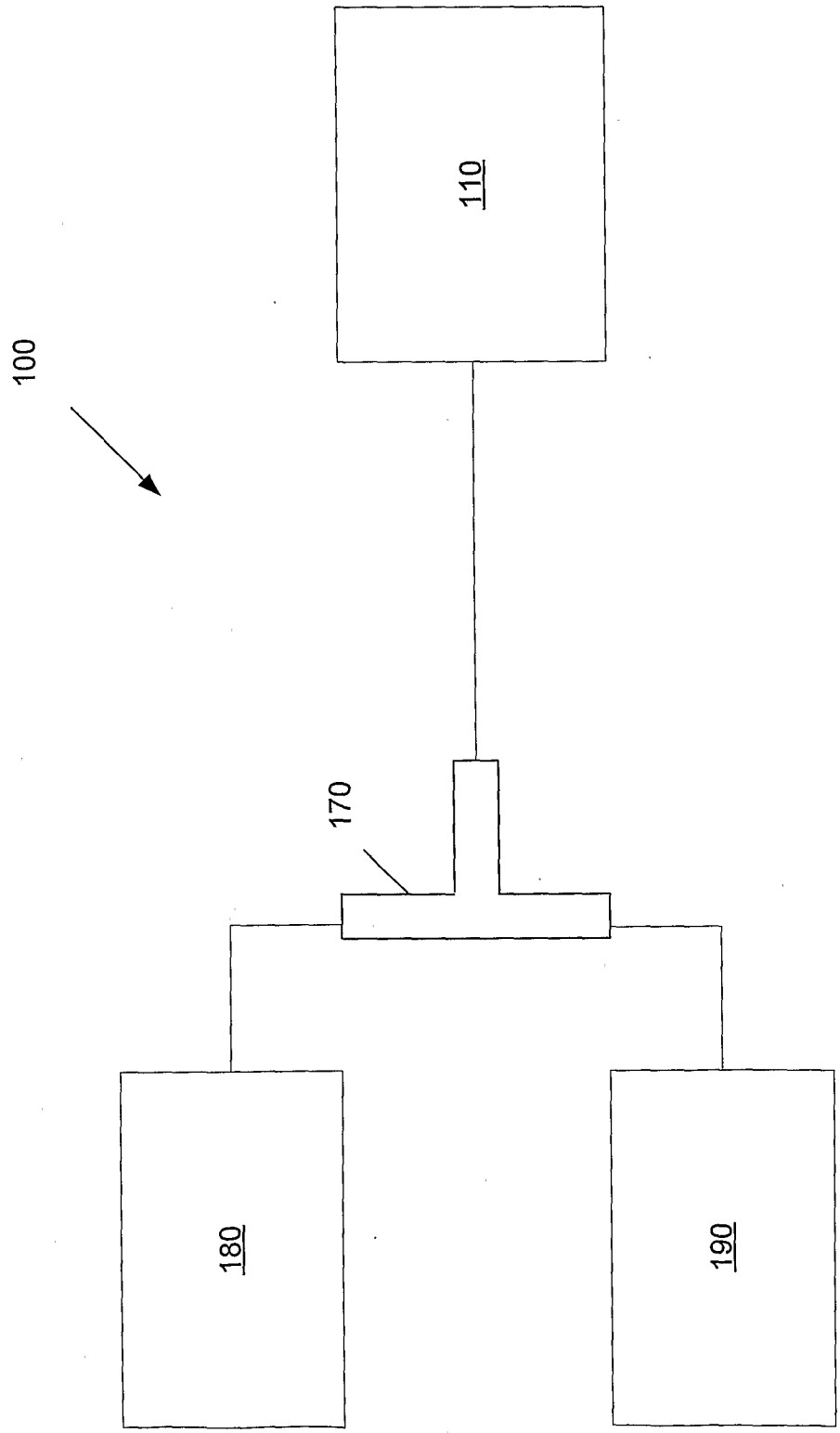


Fig. 1

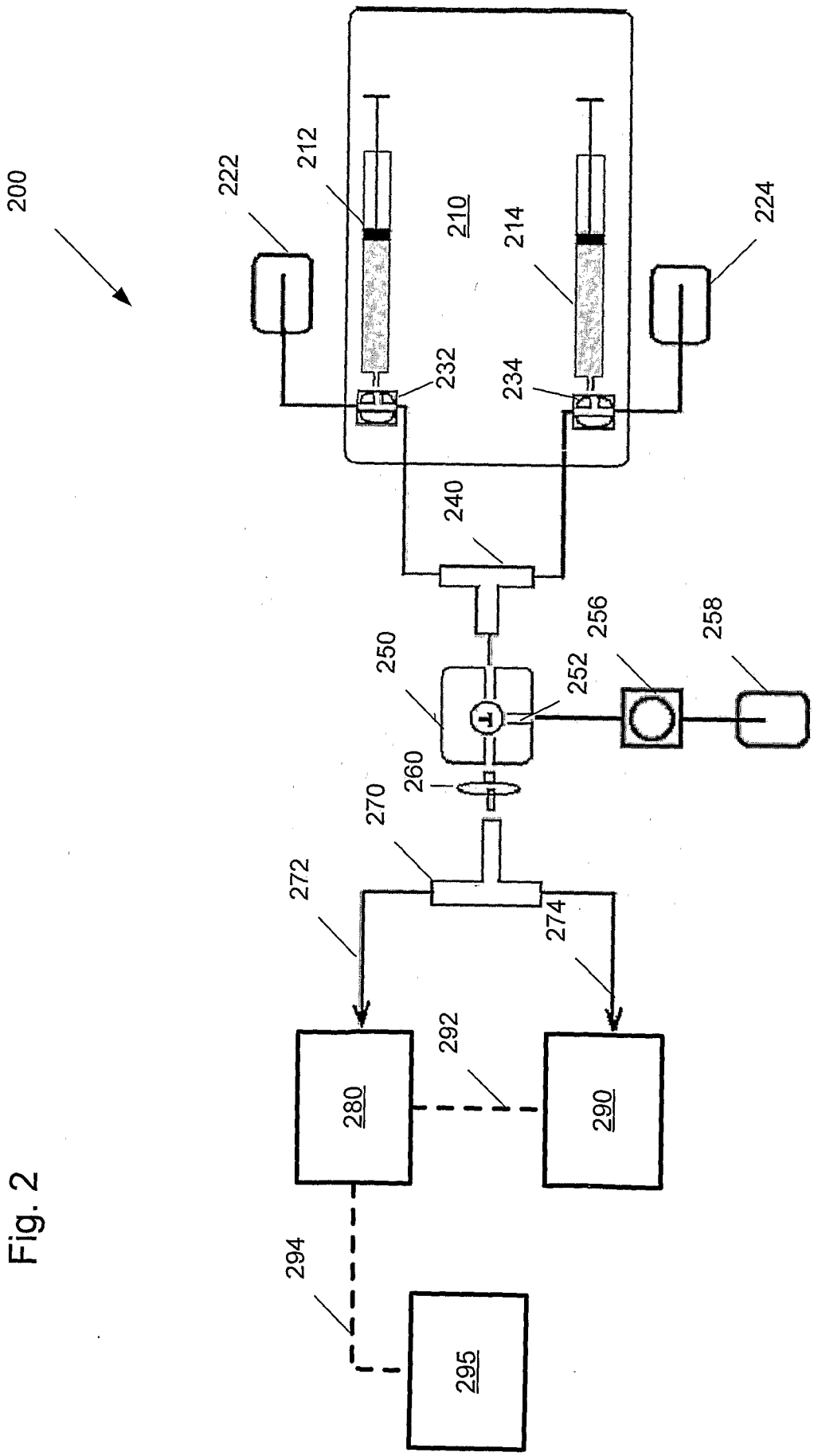

200
222
212
210
214
224
232
234
240
250
252
256
258
260
270
272
274
280
290
292
294
295

Fig. 2

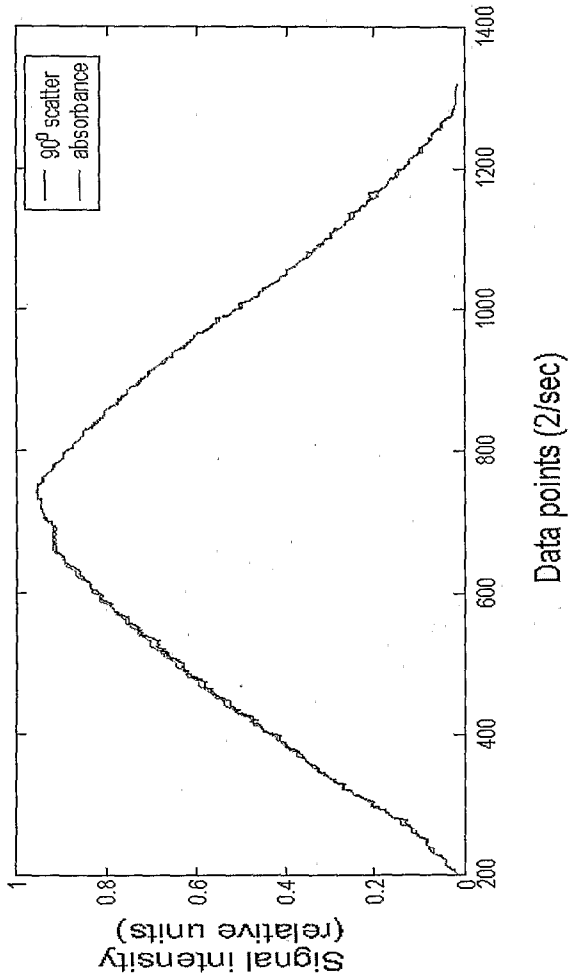

90° scatter
absorbance
Signal intensity
(relative units)
Data points (2/sec)
0
0.2
0.4
0.6
0.8
1
200
400
600
800
1000
1200
1400

Fig. 3

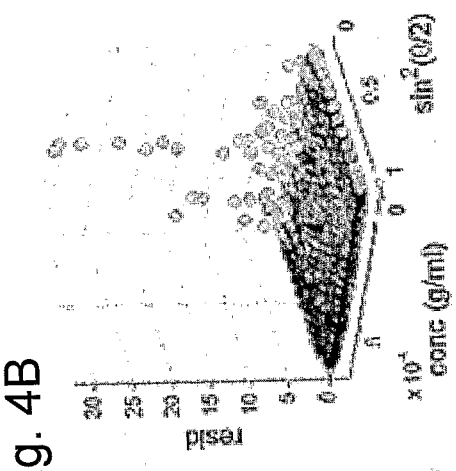


Fig. 4B

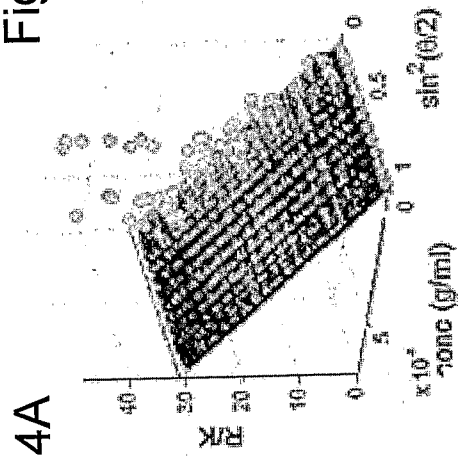


Fig. 4A

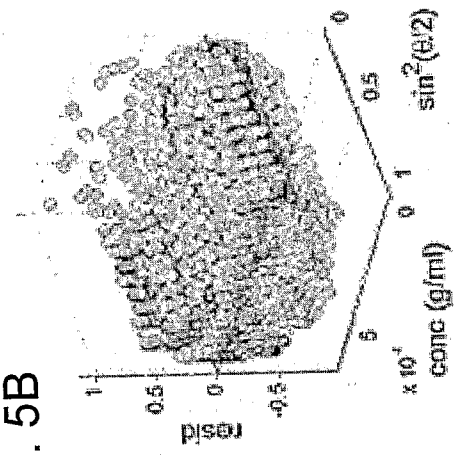


Fig. 5B

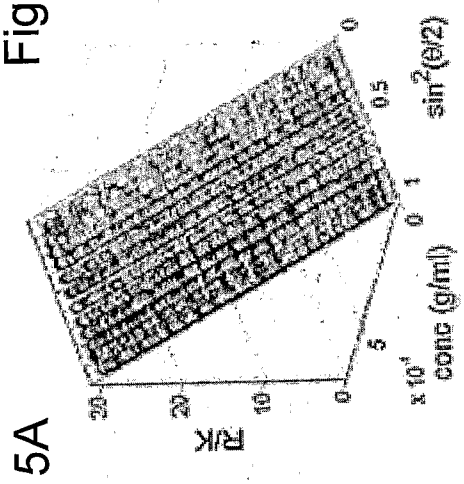


Fig. 5A

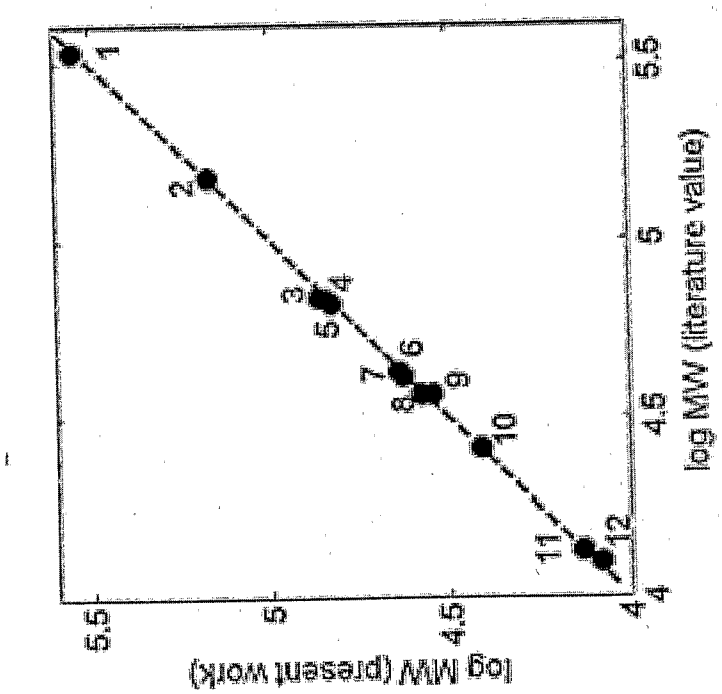

log MW (present work)
log MW (literature value)
5.5
5
4.5
4
1
2
3
4
5
6
7
8
9
10
11
12

Fig. 6

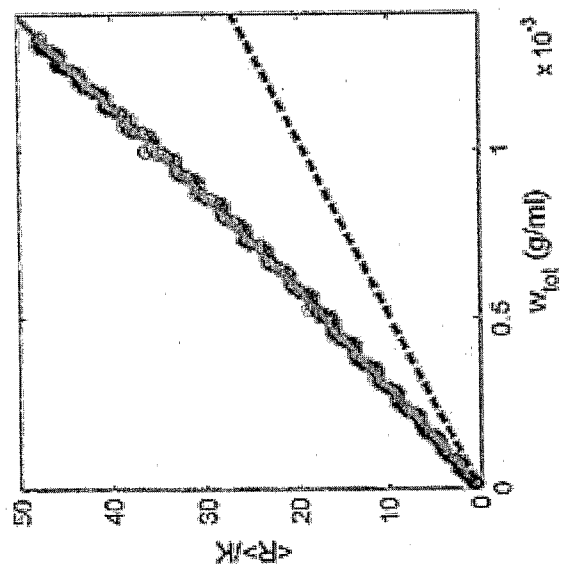


Fig. 7

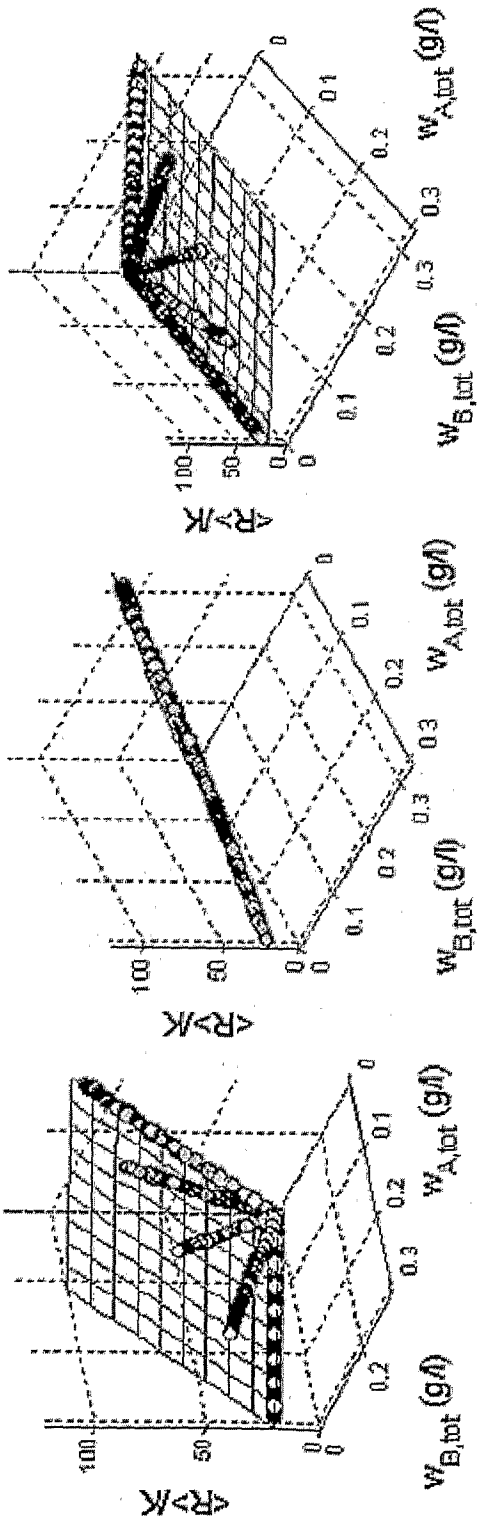


Fig. 8C

Fig. 8B

Fig. 8A

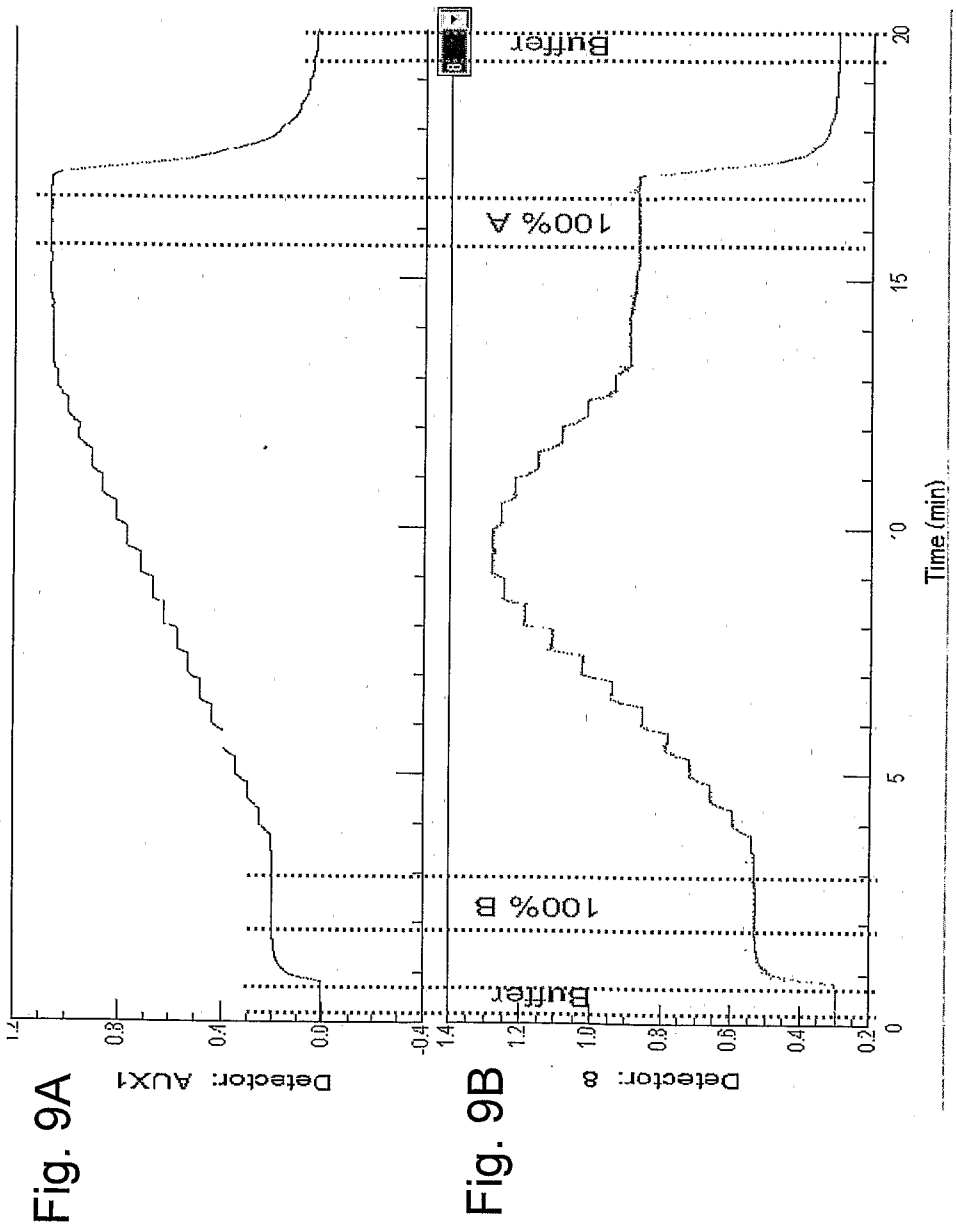


Fig. 9A

Fig. 9B

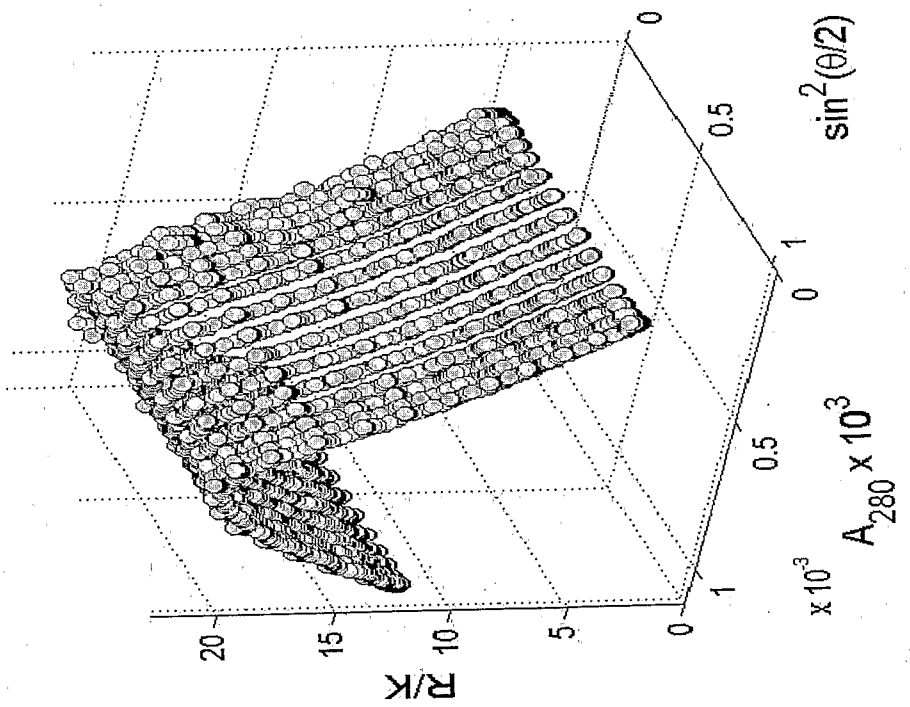


Fig. 10

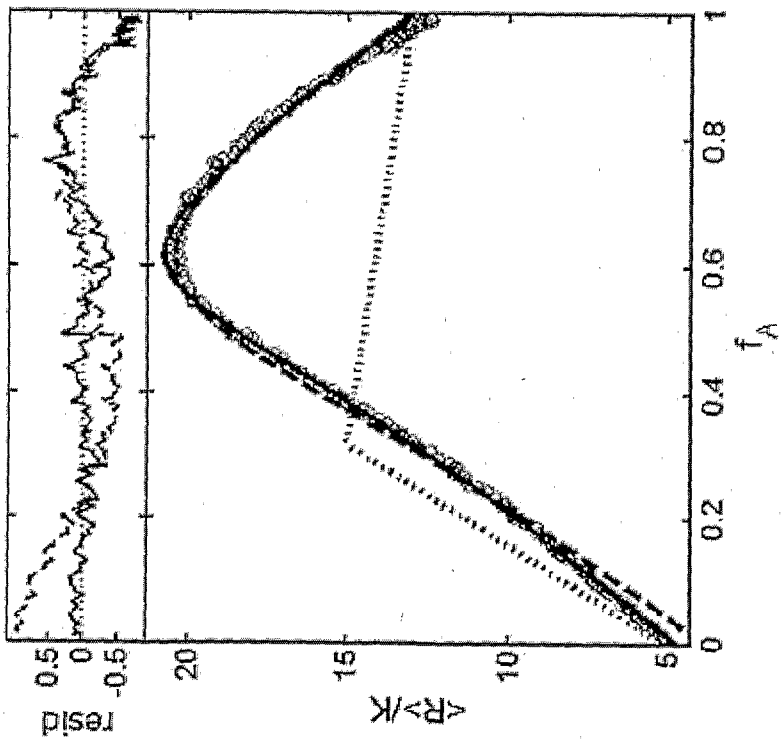


Fig. 11

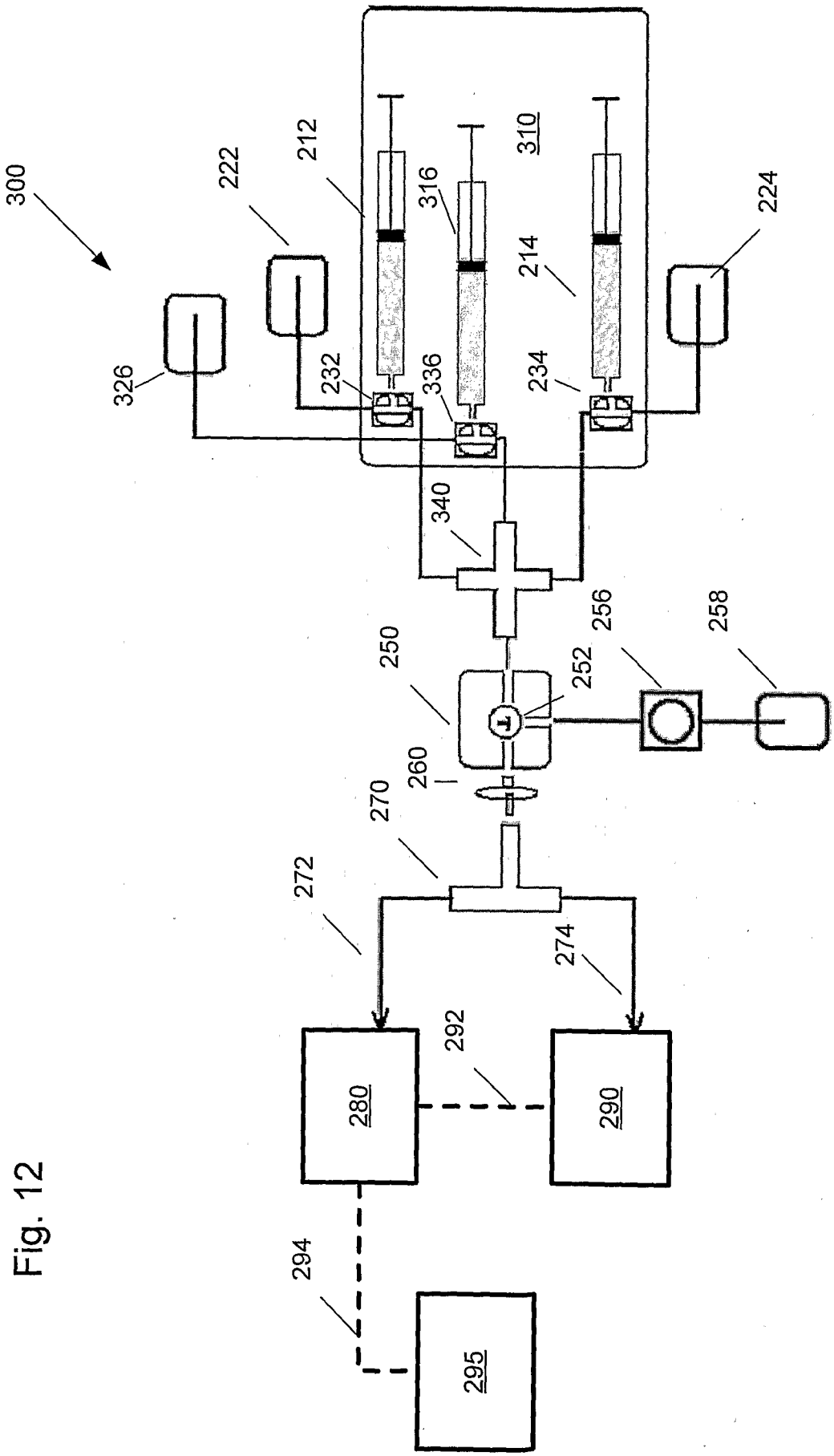

300
222
212
310
316
224
214
326
232
336
234
340
256
258
250
252
260
270
272
274
280
292
290
294
295

Fig. 12

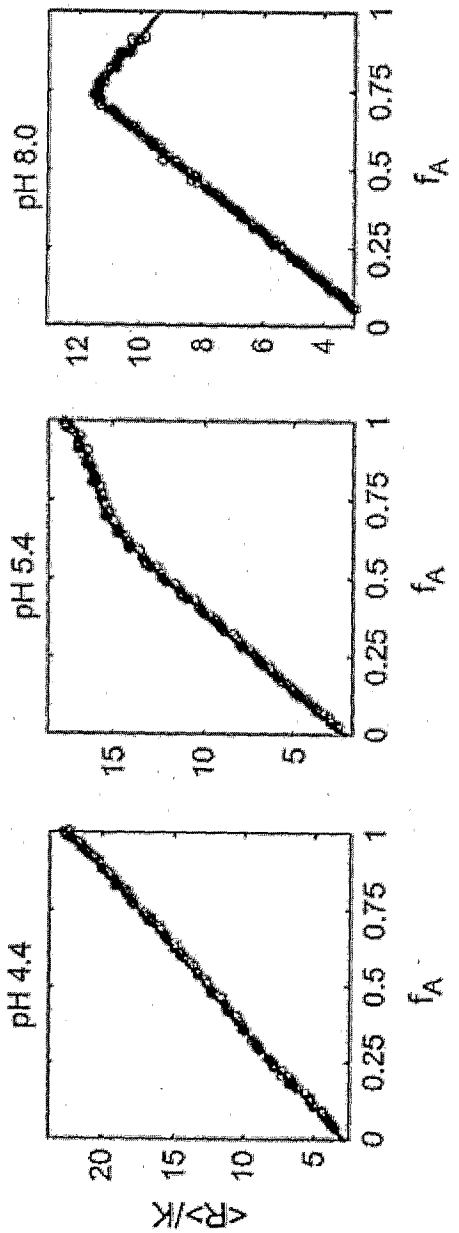


Fig. 13A

Fig. 13B

Fig. 13C

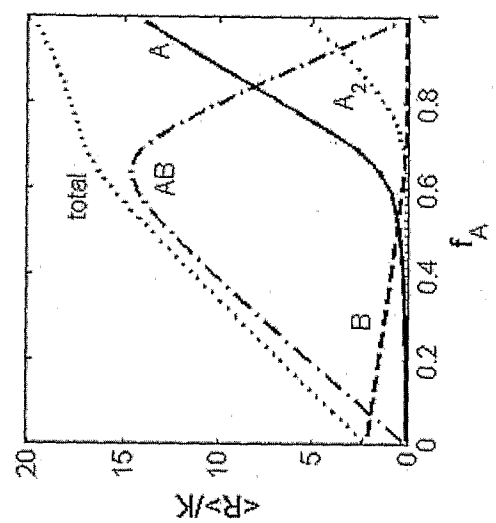


Fig. 14B

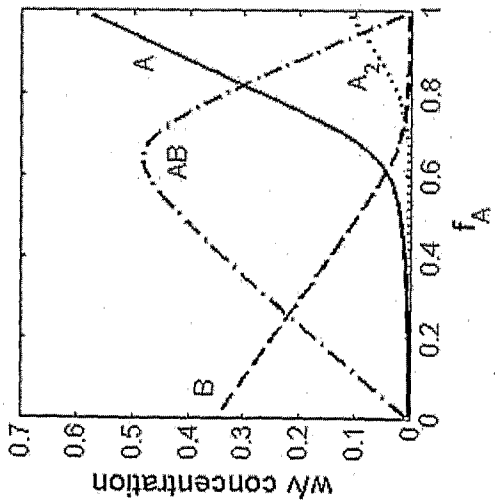


Fig. 14A

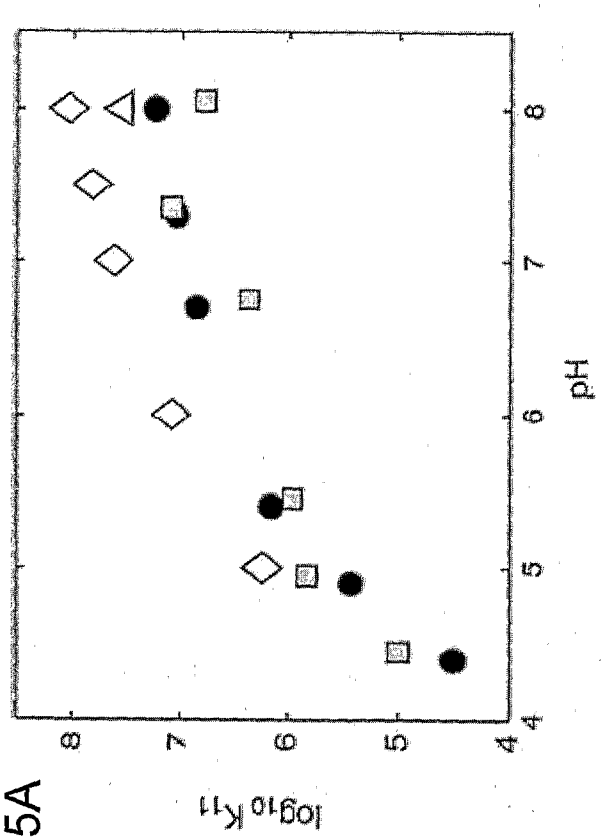


Fig. 15A

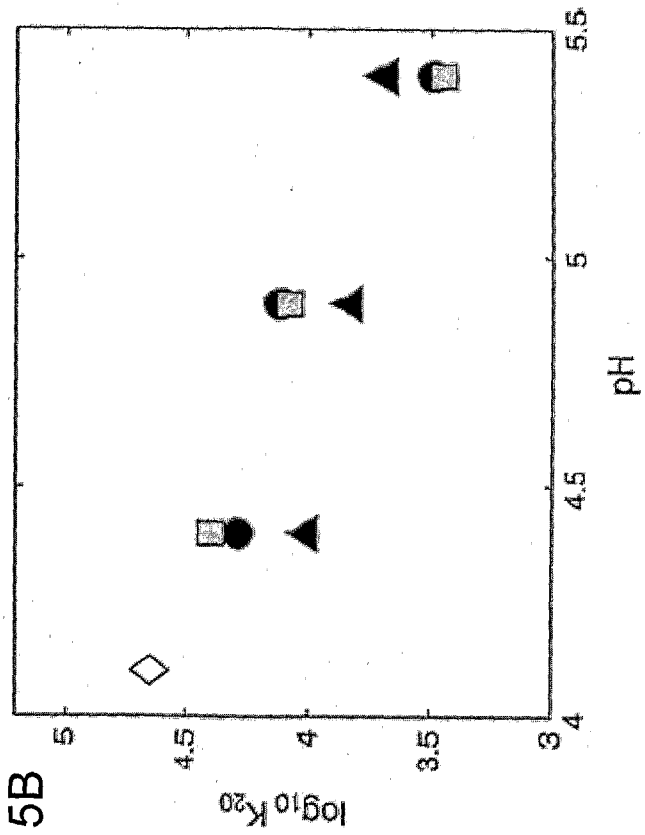


Fig. 15B

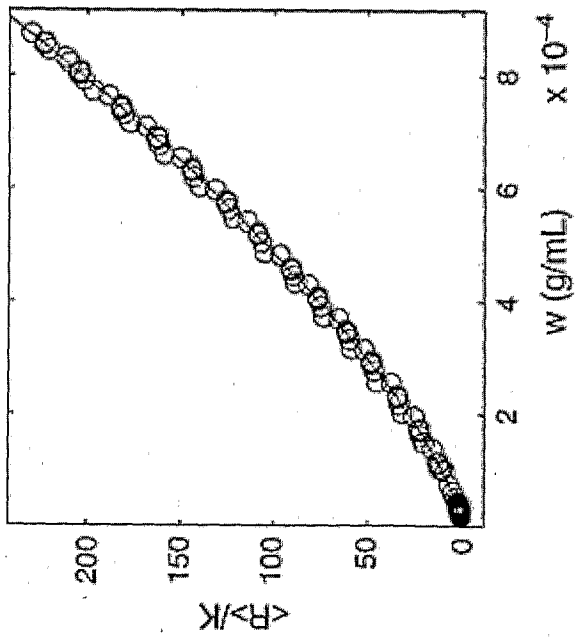


Fig. 16

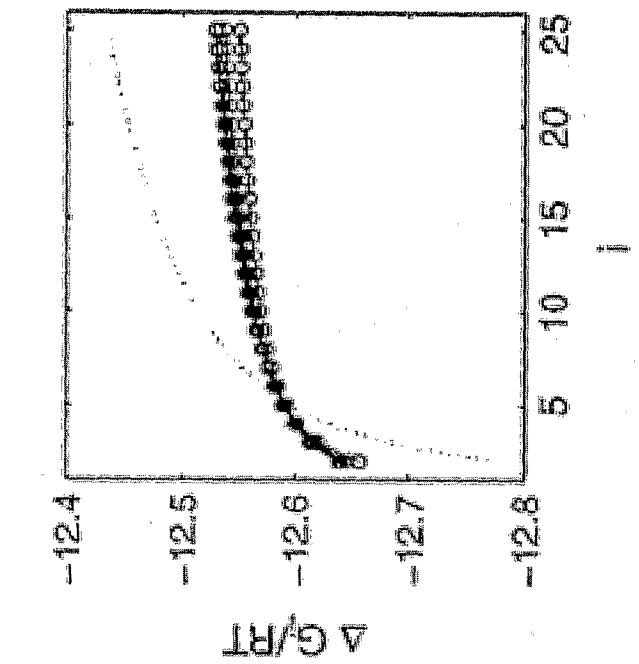


Fig. 17B

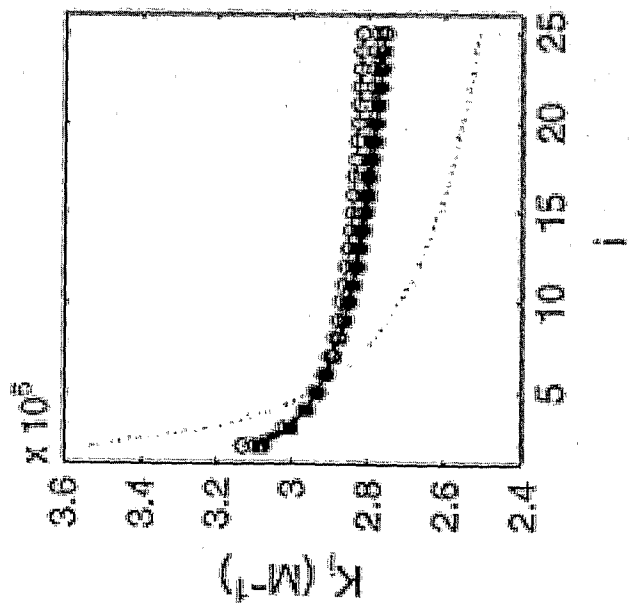


Fig. 17A

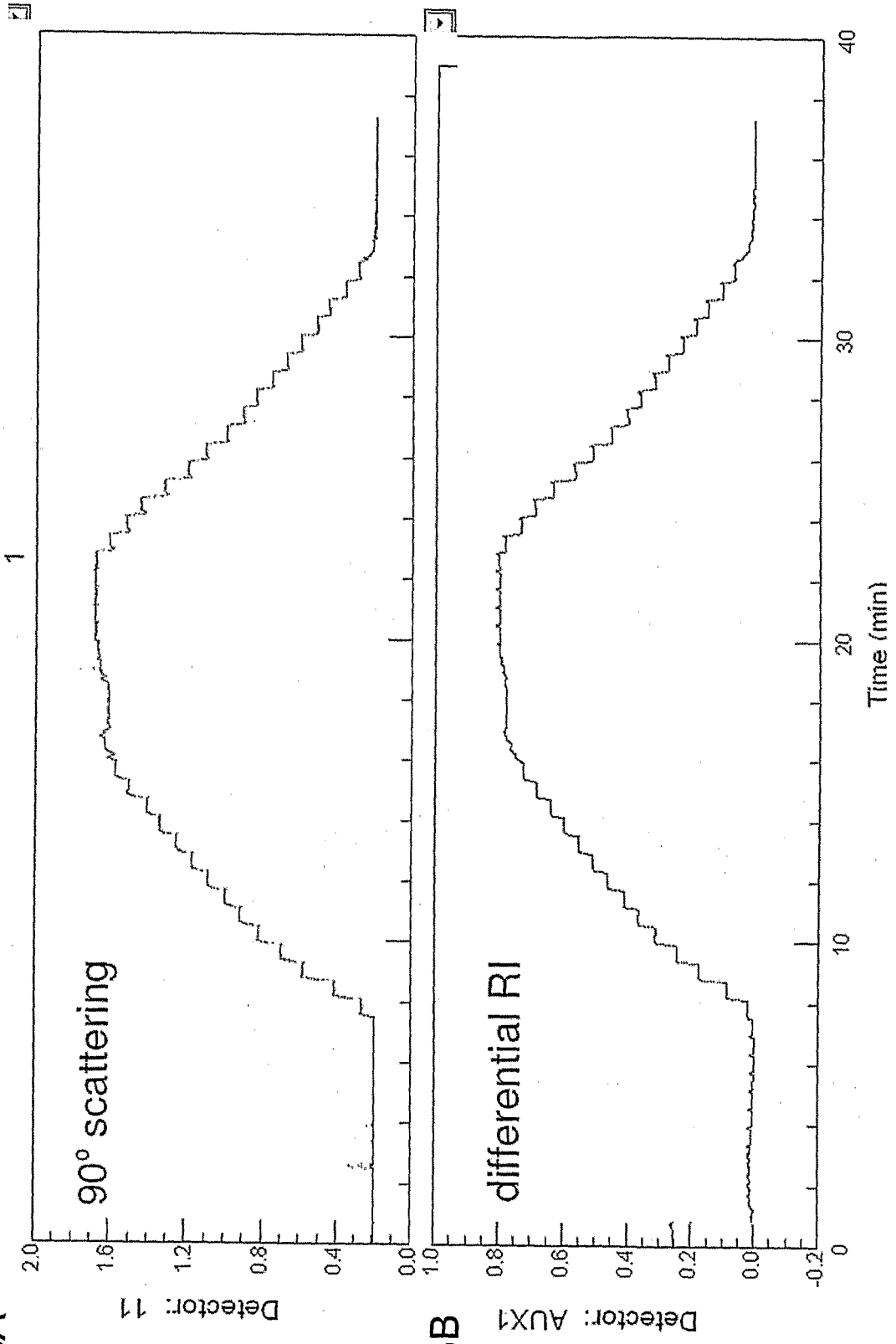


Fig. 18A

Fig. 18B

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2006/029413

A. CLASSIFICATION OF SUBJECT MATTER
INV. G01N35/10 G01N35/08 G01N21/53

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, WPI Data, BIOSIS, EMBASE, INSPEC, COMPENDEX

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ATTRI A K ET AL: "New methods for measuring macromolecular interactions in solution via static light scattering: basic methodology and application to nonassociating and self-associating proteins" ANALYTICAL BIOCHEMISTRY, ACADEMIC PRESS, NEW YORK, NY, US, vol. 337, no. 1, 1 February 2005 (2005-02-01), pages 103-110, XP004707214 ISSN: 0003-2697 figure 1 the whole document	1-19

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
23 November 2006	12/12/2006

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Timonen, Tuomo

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2006/029413

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2005/019933 A1 (ANDERSSON KARL [SE] ET AL) 27 January 2005 (2005-01-27) figures 1,10 paragraphs [0032], [0036], [0037] paragraphs [0067], [0095]	1-19
A	EP 1 510 807 A (WYATT TECHNOLOGY [US]) 2 March 2005 (2005-03-02) figure 1 paragraph [0028]	
A	US 2003/142309 A1 (KUEBLER SIGRID C [US] ET AL) 31 July 2003 (2003-07-31) paragraph [0059] paragraph [0072]	1-19
A	WO 01/57520 A2 (TRUTNAU HANS HEINRICH [DE] LEICA MICROSYSTEMS INC [US]; TRUTNAU HANS-H) 9 August 2001 (2001-08-09) the whole document	1-19
A	DE 199 07 448 A1 (SPOHN UWE [DE]; FUHRMANN BODO [DE]) 31 August 2000 (2000-08-31) figure 1	1-19
A	EP 1 308 709 A2 (MICROMASS LTD [GB]) 7 May 2003 (2003-05-07) the whole document	1-19
X,P	ARUN K. ATTRI, ALLEN P. MINTON: "Composition gradient static light scattering: A new technique for rapid detection and quantitative characterization of reversible macromolecular hetero-associations" ANALYTICAL BIOCHEMISTRY, vol. 346, no. 1, 8 September 2005 (2005-09-08), pages 132-138, XP002408695 the whole document	1-19

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2006/029413

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2005019933 A1	27-01-2005	NONE	
EP 1510807 A	02-03-2005	NONE	
US 2003142309 A1	31-07-2003	NONE	
WO 0157520 A2	09-08-2001	AT 344455 T	15-11-2006
		AU 4237401 A	14-08-2001
		DE 10005301 A1	09-08-2001
		EP 1259810 A2	27-11-2002
		US 2003143565 A1	31-07-2003
DE 19907448 A1	31-08-2000	NONE	
EP 1308709 A2	07-05-2003	CA 2411021 A1	01-05-2003
		EP 1308708 A2	07-05-2003
		GB 2383842 A	09-07-2003
		US 2003106945 A1	12-06-2003
		US 2003109061 A1	12-06-2003