Analytical Biochemistry **299,** 98–107 (2001)
doi:10.1006/abio.2001.5396, available online at http://www.idealibrary.com on IDEAL®

A New Data Analysis Method to Determine Binding Constants of Small Molecules to Proteins Using Equilibrium Analytical Ultracentrifugation with Absorption Optics[1]

Michelle Arkin* and J. D. Lear†

**Sunesis Pharmaceuticals Inc., 341 Oyster Point Boulevard, South San Francisco, California 94080; and †Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6059*

Received July 16, 2001; published online November 3, 2001

In principle, equilibrium analytical ultracentrifugation (AU) can be used to quantify the binding stoichiometry and affinity between small-molecule ligands and proteins in aqueous solution. We show here that heteromeric binding constants can be determined using a data-fitting procedure which utilizes a postfitting computation of the total amount of each component in the centrifuge cell. The method avoids overconstraining the fitting of the radial concentration profiles, but still permits unique binding constants to be determined using measurements at a single wavelength. The computational program is demonstrated by applying it to data obtained with mixtures of a 500-Da molecule and interleukin-2, a 16-kDa protein. The 1:1 binding stoichiometry and heteromeric dissociation constants (K_{ab}) determined from centrifuge data at two different wavelengths are within the 4–9 μM range independently determined from a functional assay. Values for K_{ab} have been obtained for ligands with affinities as weak as 500 μM. This AU method is applicable to compounds with significant UV absorbance (~0.2) at concentrations within ~5- to 10-fold of their K_{ab}. The method, which has been incorporated into a user procedure for IgorPro (Wavemetrics, Oswego, OR), is included as supplementary material.

Sedimentation equilibrium analysis by analytical ultracentrifugation (AU)[2] offers a powerful and versatile method for studying biologically important molecules in aqueous solutions (1–3). The method in general relies on the analysis of concentration gradients produced by the differential sedimentation of species with different molecular weights. Modern, computerized instrumentation and data analysis have so increased the efficiency of this method that it is now applied routinely to the determination of protein–protein and protein–DNA binding equilibrium constants (e.g., 4–6). However, application of equilibrium AU to the binding of small molecules to larger proteins appears to have been relatively neglected. In part, this is due to the perceived difficulty in reliably detecting the small change in molecular weight of the large molecule upon binding of a much smaller one.

This problem is easily overcome once it is recognized that the smaller molecule experiences a very large molecular weight increase when it binds. Binding of the small-molecule ligand is readily visualized when the compound absorbs at a wavelength where the protein does not. Figure 1 shows such a situation schematically. A small molecule, detected at 320 nm, does not sediment when centrifuged at 25K rpm in the absence of protein. When a binding protein is added, the apparent mass of the small molecule increases, yielding a large concentration gradient at 320 nm and an offset due to the unbound small molecule. As shown in Fig. 1, the observed gradient can be deconvolved into its various components—unbound compound, unbound pro-

[1] Supplementary data for this article are available on IDEAL (http://www.idealibrary.com).

[2] Abbreviations used: AU, analytical ultracentrifugation; IC_{50},inhibition constant (ligand concentration at 50% inhibition); K_d (K_{ab}), dissociation constant of *ab* heterodimer; DMSO, dimethyl sulfoxide; RMS, root mean square; χ_{mat}, root mean square deviation between calculated and known component concentrations; IL-2, interleukin-2; IL-2R, IL-2 receptor; PBS, phosphate-buffered saline; SPA, scintillation proximity assay.

0003-2697/01 $35.00

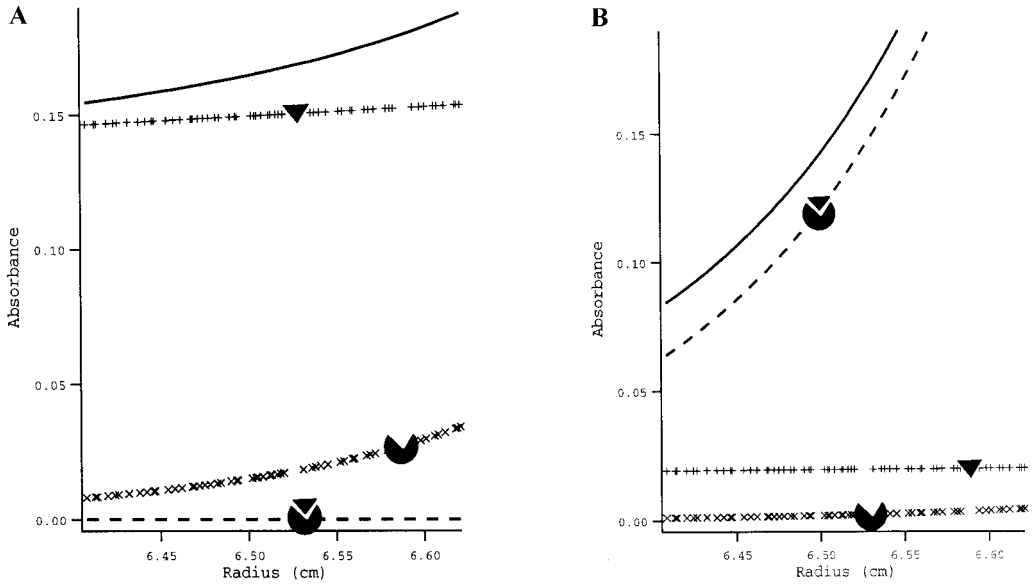
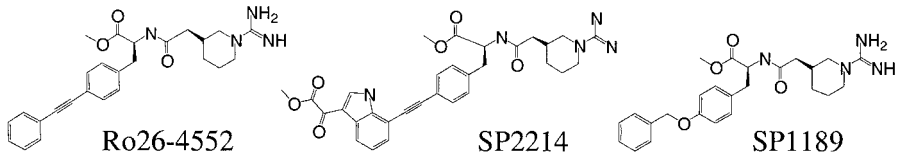


FIG. 1. Simulation showing equilibrium sedimentation radial concentration profiles for 5 μM small molecule (small triangle) with MW = 500 Da, extinction coefficient (ϵ) = 25,000 M^{-1} cm^{-1} in the presence of 5 μM of a larger molecule (15 kDa, ϵ = 3000 M^{-1} cm^{-1}; large, indented shape). (A) Compound does not bind protein. (B) Compound binds to protein with 1:1 stoichiometry and a dissociation constant of 10^{-7} M. The solid lines show the observed absorption signal, the dashed lines show the contribution of the 1:1 complex (combined shapes). Contributions from unbound monomers are shown by +'s (low Mw component) and ×'s (high MW).

tein, and the compound/protein complex. Qualitatively, one can determine that a compound is binding simply by observing a gradient in the presence of protein. Such an analysis has been used to rank relative binding affinities and to approximate the binding stoichiometry (7–9). This experiment would be much more valuable, however, if binding measurements were (a) quantifiable and (b) applicable to wavelengths where both compound and protein are observed.

There are a small number of published methods designed explicitly for the determination of heteromeric dissociation constants (K_{ab}) by sedimentation equilibrium (2, 10–12). In addition, a rigorous thermodynamic treatment including activity coefficients has been shown to be applicable to heteromeric systems (13–15). In every case, however, quantifying the binding of small-molecule ligands to larger proteins by equilibrium AU requires a quantitative consideration of component mass balance. This is because the equations which define sedimentation equilibrium rely only on the shape of gradient curves and not on the concentration of materials, some of which have sedimented to the cell bottom. Published methods have used different approaches to constrain the component concentrations in order to maintain mass balance and to deconvolve the separate contributions from the protein and ligand at wavelengths where both components absorb. For example, the "multiwavelength scan method" has been used by Kim *et al.* to quantify the binding of a 13-bp DNA fragment to the DNA-binding domain of heat shock transcription factor (16). In this work, Kim *et al.* took advantage of the fact that DNA and proteins have readily distinguishable UV absorption spectra. Solutions of DNA and protein were monitored at several wavelengths and the contribution from each component was deduced by simultaneously solving the equilibrium sedimentation equations for each data set. Earlier, Lewis developed the method of "implicit constraints" (mentioned in 10) which computes the integrated concentrations of the two different molecules to provide the additional data constraint necessary for a unique fit to the data. Programs developed by Minton (TWOCOMP, discussed below) (11) also employ this constraint. Minton recommended using both multiple-wavelength scanning and loading concentration constraints to reduce unwanted parameter correlation in data fitting. Although loading concentration constraints alone should be sufficient to obtain mass balance, errors in initial loading concentrations and cell volumes can seriously distort data fitting.

We set out to develop AU as a quantitative binding assay for small-molecule antagonists of protein–protein interactions. Protein–protein interactions govern a host of biological processes and therefore represent an important class of therapeutic targets. However, traditional methods of small-molecule drug discovery have often yielded compounds with nonspecific binding sites and ill-defined mechanisms. It is therefore important to develop simple methods for verifying affinity and stoichiometry of binding. For our method-develop-



SCHEME 1

ment studies, we use the interleukin-2 (IL-2)–IL-2 receptor (IL-2R) system as a model. IL-2 represents an important drug target because binding of IL-2 to IL-2R triggers proliferation of T-helper cells, a cardinal event in autoimmune disease and graft rejection. Furthermore, Tilley *et al.* (17) have published the discovery and characterization of a synthetic small-molecule ligand (Ro26-4550; Scheme 1) that binds to IL-2 and inhibits the interaction of IL-2 with the α chain of IL-2R. Arkin and co-workers have further characterized the binding of Ro26-4550 to IL-2 by X-ray crystallography and have used this system to validate and compare biophysical methods such as NMR, calorimetry, and surface plasmon resonance (18 and unpublished results). Collectively, these experiments show that Ro26-4550 binds with 1:1 stoichiometry to IL-2 at the IL-2Rα binding site with a $K_{ab} \sim 5$ μM.

In this paper, we apply a novel AU data analysis method to determine the binding constants of small-molecule ligands to IL-2. Our method, named "HetFitter," differs from those previously described in that it utilizes a postfitting computation of loading concentrations to provide a secondary criterion for evaluating data fits. This approach avoids both parameter ambiguity and fit distortion and allows rational choices to be made among different models giving nearly equivalent fits to radial concentration profiles. We demonstrate the use of HetFitter by determining the binding characteristics of SP2214 and SP1189 (18), Ro26-4550 derivatives synthesized in our laboratory. These compounds represent two typical experimental conditions; the absorbance spectrum for SP2214 extends to wavelengths where IL-2 does not absorb, whereas SP1189 has a weak absorbance spectrum which overlaps that of IL-2. By applying our data analysis method to equilibrium AU data determined at a single wavelength, we are able to determine unique binding constants between these compounds and IL-2. The dissociation constant obtained for IL-2/SP2214 is nearly the same at 320 nm, where the protein is not detectable, and at 290 nm, where compound and protein are both observed. Moreover, for a series of compounds, the rank ordering of K_{ab}'s determined by our method agrees closely with inhibition constants obtained from functional assays. These measurements use about 100 μg of protein and require only that the compound show significant UV absorbance at concentrations within 5- to 10-fold of its binding affinity. AU thus provides an alternative to methods such as isothermal calorimetry and NMR, both of which require large amounts of material and high compound solubility.

MATERIALS AND METHODS

AU measurements employ the Beckman XLA, which measures the optical density of a solution at fixed, chosen wavelengths along the radius of a centrifuging cell. The resulting radial profile represents the total concentration of optically absorbing species in the cell. This signal, $S(r)$, is the sum of the absorbances, related to their respective concentrations through Beer's law, of each of the contributing species. For associating components (a and b), the equilibrium concentrations of the associated species (ab) at each radial position can be related to the component monomer concentrations ($a + b$) by suitably defined equilibrium constants (K_{ab}). We employ the following equation to describe the radial concentration distribution associated with equilibrium sedimentation in centrifugation:

$$S(r) = \sum_n \frac{1.2 n_a \epsilon_a [C_a]^{n_a} \exp\{n_a \chi M_a\}}{K_{na}} + \sum_n \frac{1.2 n_b \epsilon_b [C_b]^{n_b} \exp\{n_b \chi M_b\}}{K_{nb}} + \sum_n 1.2 \frac{(n_a \epsilon_{ab} + n_b \epsilon_b)\left[\frac{C_a}{\epsilon_a}\right]^{n_a}\left[\frac{C_b}{\epsilon_b}\right]^{n_b} \times \exp\{\chi(n_a M_a + n_b M_b)\}}{K_{nab}} + baseline, \quad [1]$$

where $S(r)$ is signal due to all sedimenting species at radial position r; $\epsilon_{a,b}$ is the monomer molar extinction coefficient of a, b; ϵ_{ab} is the molar extinction coefficient component a when complexed with b; $C_{a,b}$ is the molar concentration at r_o of monomer of molecular weight $M_{a,b}$; $\chi = ((1-\bar{\nu}\rho)\omega^2)/2RT(r^2 - r_o^2)$; r_o is the arbitrary fixed radius reference; $n_{a,b}$ is the number of a,b monomers in particular association state being considered; $K_{na,nb,nab}$ is the dissociation constant in units of (e.g.) Molar $^{(na-1)}$for a,b,ab species; $\bar{\nu}$ is the partial specific volume (cm^3/g) of all sedimenting species (assumed equal); ρ is the density of supporting buffer (g/cm^3); ω is

the angular velocity of rotor (radians/s); $M_{a,b}$ is the molecular weight of monomer a,b species (g/mol); R is the ideal gas constant (8.315×10^7 ergs K^{-1} mol^{-1}); and T is the temperature (K).

This equation, derived from elementary principles of equilibrium sedimentation (e.g., 2, 19, 20) has three terms, representing the gradient at sedimentation equilibrium for the protein, ligand, and complex, respectively. Each term allows for oligomerization, given by n. The absorbance of each species is given by Beer's law, $1.2 * n * \epsilon * C_i$ (where ϵ is the monomer extinction coefficient, 1.2 is the pathlength in the centrifuge cell, and C_i is the volume concentration of species i). For simplicity, the partial specific volume is assumed to be the same for all species. This assumption could readily be relaxed if components of widely different partial specific volumes were to be studied. In most cases, an adjustment of the assumed molecular weight of one of the components should suffice to compensate for buoyancy differences. The partial specific volume for IL-2 and density of buffer were calculated using the software Sednterp (21).

A function using this equation was implemented within an IgorPro (Wavemetrics, Oswego, OR) User Procedure (called HetFitter; see supplementary material) and used to analyze SE curves via nonlinear least square data fitting using IgorPro's Levenberg–Marquardt algorithm. After fitting experimental (or simulated) $S(r)$ curves, another, separate function within the procedure uses the best-fit parameters to integrate the equilibrium concentrations of each species over the entire volume of solution in the centrifuge cell. From these integrated concentrations, the function computes the total concentration of each component (a and b) and compares the result with the known (user-input) concentrations.

The computation of concentrations is indispensable in fitting two component equilibria because sedimentation equilibrium calculations are based entirely on concentration *gradients.* Without constraints on component concentrations, values obtained by curve fitting alone can easily be inconsistent with values required for material balance over the cell. In the TWOCOMP programs developed by Minton (11), total concentration of each component in the cell is used as a constraint in curve fitting. We have found, however, that small errors in concentrations can so bias the curve fit that calculated radial concentration profiles deviate unacceptably from the observed data. To avoid this problem, our method fits radial concentration profiles with freely varying component concentrations, but over a range of fixed binding parameters. Each round of fitting at a fixed dissociation constant produces a best-fit radial concentration profile and a value of the predicted integrated concentration. Using the summed root-mean-square (RMS) deviation between the calculated and known component concentrations (χ_{mat}) provides a systematic way to distinguish among differing parameter sets giving equally good fits to radial concentration profiles.

To provide a comparison of our fitting method with those utilizing strict material balance constraints, we downloaded Minton's TWOCOMP software package from the public site ftp://alpha.bbri.org/rasmb/spin/ms_dos/twocomp-minton/. Solution concentration (g/liter), extinction coefficient (1.2 liter/g), and equilibrium constant $(g/liter)^{(1-n)}$ units in that program (described in 11) were related to those used in our program by application of appropriate, molecular-weight-dependent conversion factors. The module "2CS1Fit" was employed since no more than three data sets needed to be fit simultaneously. The "fcomp1" and "fcomp2" parameters (defined as the fractions of component competent to undergo further association) were held constant at unity for initial fitting and then allowed to vary freely to refine curve fits. Component loading concentrations, sample meniscus, and cell bottom positions were manually input to be the same as those used in HetFitter.

Sample preparation. Compounds were synthesized at Sunesis Pharmaceuticals (18). Dimethyl sulfoxide (DMSO) and buffer salts were purchased from Sigma. Water was purified by Milli-Q filtration system (Millipore). Stock solutions were made by dissolving weighed compounds into DMSO. Solutions for absorbance and AU measurements were prepared by diluting DMSO stocks in phosphate-buffered saline to a final DMSO concentration of 1% (v/v). IL-2 was cloned and expressed in *E. coli* as previously described (22, 23).

Absorbance spectra. UV-visible absorbance spectra were measured with an Agilent 8453A spectrometer, using microcuvettes (Hellma) with 70 μl volume and 1 cm pathlength. Extinction coefficients were determined from the absorbance of PBS/1% DMSO solutions according to Beer's Law. These spectra and calculated extinction coefficients are used only to determine appropriate wavelengths and concentrations for AU studies. Extinction coefficients calculated from the spectrometer differed from those calculated from the AU data by 0–50%, depending on the wavelength measured and the intensity of the XLA lamp. Typical variances are within 25%.

AU measurements. Equilibrium sedimentation experiments utilized six-sector cells in an eight-cell rotor (Beckman). Samples were set up in one of two ways. In the first, compounds (40 μM) were added to 8, 40, or 200 μM IL-2 in sectors A–C of one cell. Forty micromolars of each compound were monitored in the same experiment in another cell. It may be preferable to have the compound in the same cell as the compound/protein samples, however, since we have found that

inaccuracies in the absorbance measurement between cells can complicate the determination of extinction coefficients, particularly when the absorbance spectra contain sharp peaks. Several experiments (including Figs. 3 and 4) were therefore set up with compound in sector A, protein in sector B, and compound plus protein in sector C. While fewer compounds can be analyzed by this approach, we recommend it for obtaining more accurate extinction coefficients and therefore more accurate dissociation constants. Experiments with SP2214 used 25 μM compound and varying protein concentrations (5, 25, 100 μM). Experiments with SP1189 used 80 μM compound, with protein concentrations of 40 and 80 μM.

Equilibrium AU experiments were done as follows: at a rotor speed of 3K rpm, wavelength scans (spectra) were measured for each sample, followed by radial scans. The extinction coefficient used for fitting was determined from these 3K radial scans. The centrifuge was then set at 20K or 25K rpm and run for 20–24 h before radial scans were collected at each wavelength of interest. Previous studies at multiple speeds showed that our preparation of IL-2 behaves as a monomer at these concentrations. The 20–25K rpm scans were used for K_{ab} determination by HetFitter and TWOCOMP. Finally, the centrifuge was set at 50K rpm and run for 20–24 h before radial scans were collected at the same wavelengths. The absorbance at the meniscus in the protein-only sectors was used to determine an appropriate baseline value to use in the fit. These absorbance values were sufficiently close to zero ($<$0.01 OD) that, to facilitate comparisons with the TWOCOMP program, we assumed zero values for curve fitting (see "Notes on Accuracy and Reliability" below). The 50K rpm data are also useful for qualitatively assessing compound binding by observing the absorbance at the meniscus in sectors containing both compound and protein.

Scintillation proximity assay (SPA). The inhibitory constant (IC_{50}) is defined as the ligand concentration at 50% inhibition of the IL-2–IL-2Rα complex. The IC_{50} values for all compounds were determined by SPA using 12 threefold serial dilutions of compound. Each well of a 12-point titration included compound, 10 nM IL-2 labeled with tritiated propionic acid, and scintillant-containing beads labeled with streptavidin and saturated with biotinylated IL-2Rα (0.7 mg/ml beads; streptavidin beads from Amersham), in a final volume of 100 μl Superblock with 1% DMSO. Scintillation due to IL-2 bound to IL-2Rα was read on a Trilux scintillation counter (Applied Biosystems). IC_{50} values were determined from curves of scintillation counts versus compound concentration using a nonlinear regression analysis.

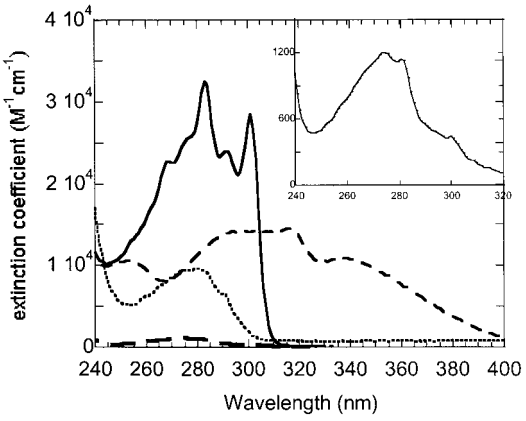


FIG. 2. Absorbance spectra for IL-2 and small-molecule ligands. The main graph shows the optical absorption extinction coefficients (M^{-1} cm^{-1}) versus wavelength (nm) for Ro26-4550 (dashed line), SP2214 (solid line), SP1189 (long-dashed line), and IL-2 (dotted line). Inset shows the absorbance spectrum (extinction coefficient versus wavelength) for SP1189.

EXPERIMENTAL RESULTS

IL-2 Binding to Small-Molecule Ligands

Two experimental situations can be studied by equilibrium AU. In the first, the small molecule is detected at wavelengths where the protein does not absorb. In the second, both protein and small molecule are detected, and the AU signal is the sum of these components. To determine whether HetFitter was applicable in both situations, we analyzed SP2214 binding to IL-2 at 330 nm, where only the small molecule significantly absorbs light, and 290 nm, where both molecules contribute to the measured absorption signal (Fig. 2). Figure 3 shows the measured sedimentation equilibrium curves ($S(r)$'s) at a rotor speed of 25K rpm for 25 μM SP2214 bound to increasing concentrations of IL-2 (8, 25, 100 μM). Figure 3A shows data at 290 nm; Fig. 3B gives 330-nm data.

As described under Materials and Methods, $S(r)$ curves are fit globally by HetFitter in two steps. First, a dissociation constant (K_{ab}) is assumed and curves are fit to Eq. [1]. The concentrations of reagents are allowed to vary freely during this fitting procedure. In the second step, the total amounts of protein and ligand are calculated from the concentrations determined by the curve fit. The process is then repeated for different guesses of pK_{ab} and the calculated concentrations are compared to the user-entered concentrations. This is done systematically by plotting the RMS difference between the calculated and actual component concentrations (χ_{mat}) versus the negative $\log_{10}$ of the heterodimer dissociation constant in molar units (pK_{ab}). The minimum point in the χ_{mat} vs pK_{ab} curve determines the best-fit value for the dissociation constant of the protein–ligand interaction. The value of χ_{mat} at the

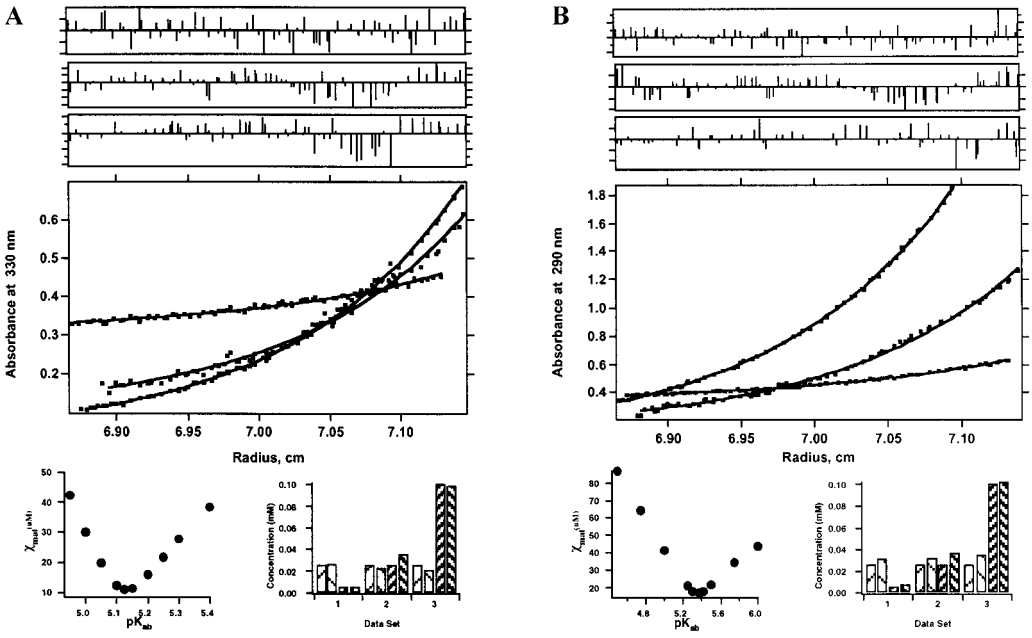


FIG. 3. Results from fitting SP2214/IL-2 AU data with HetFitter. (A) 330-nm data. (B) 290-nm data. Both A and B represent the same SP2214/IL-2 mixtures at equilibrium at a rotor speed of 25K rpm. For SP2214, the following data were input by the user: MW = 572 Da, ϵ_a = 12,500 M^{-1} cm^{-1} at both 290 and 330 nm. For IL-2, user input data included: MW = 16,000 Da, ϵ_b = 5000 M^{-1} cm^{-1} at 290 nm, ϵ_b = 67 $M^{-1}cm^{-1}$ at 330 nm. For all data sets, ν = 0.753 ml/g and ρ = 1.002 g/ml. Loading concentrations were 25 μM for SP2214 (all data sets) and 100, 25, and 5 μM for IL-2 (sets 1, 2, and 3, respectively). Cell meniscus positions were fixed to the first point used in the data set, and bottom position was determined graphically based on loss of light intensity at the chamber boundary. For both A and B, top panels show the fit residuals at the optimum pK_{ab} for sets 1, 2, and 3. Center panels show the raw data points and optimum pK_{ab} fit lines. (Bottom panels) Left, the RMS deviations of computed concentrations (χ_{mat}) versus assumed pK_{ab}; and right, histograms comparing actual (downward left hatching) and computed (downward right-hatching) concentrations for components A (light) and B (dark) for each set obtained with fitting at optimum pK_{ab} value.

minimum indicates how well the experimental data could be matched to the user-input concentrations and extinction coefficients.

The results for SP2214 binding to IL-2 are consistent with K_{ab} determined by other methods. The 7.5 μM K_{ab} calculated from the data where only the ligand is detected (330 nm) is reasonably close to the 4.2 μM K_{ab} calculated from the 290-nm data where both ligand and compound are observed. These values compare well to the IC_{50} values (4–9 μM, three measurements) determined by inhibition assays in which IL-2 is competed from IL-2Rα by SP2214.

We also measured the K_d of SP1189 bound to IL-2. The UV spectrum of SP1189 is not favorable for monitoring binding by absorbance; the maximum extinction coefficient for SP1189 is only 1250 M^{-1} cm^{-1} at 270 nm, a wavelength where IL-2 has an extinction coefficient close to 10,000 M^{-1} cm^{-1} (Fig. 2). Nevertheless, we are able to detect binding of SP1189 to IL-2 by raising the concentration of compound to 80 μM in the presence of 40 or 80 μM IL-2. SPA measurements yield an IC_{50} between 30 and 90 μM for this interaction. Similarly, two separate AU experiments give K_d values of 30 and 100 μM (Fig. 4), with χ_{mat} values of 0.030 and 0.018, respectively. For both trials, χ_{mat} vs pK_{ab} curves have a well-defined minimum; both curve fits and material balances are within acceptable limits. HetFitter is therefore able to determine reasonable values for K_d even when the compound absorbance is weaker than that of the protein.

We have observed a close correlation between K_{ab} and IC_{50} values for a series of small-molecule ligands bound to IL-2 (Fig. 5). For IC_{50} values between 5 and 500 μM, K_{ab} values determined by AU are within 2-fold of the measured inhibition constants. The experiment-to-experiment variation in both SPA and AU is ~2-fold; the correlation is therefore about as accurate as the measurements themselves. Notably, these compounds vary not only in IC_{50} values but also in absorbance spectra. Based on these data, we expect HetFitter to be applicable to compound/protein interactions with K_{ab}'s in the micromolar range, provided that the small molecules show measurable absorbance (0.2) at concentrations within approximately 5- to 10-fold of their measured IC_{50}.

Comparison of HetFitter and TWOCOMP

SP2214/IL-2 binding data were fit with the program "2C1SFIT" in the TWOCOMP package using the same

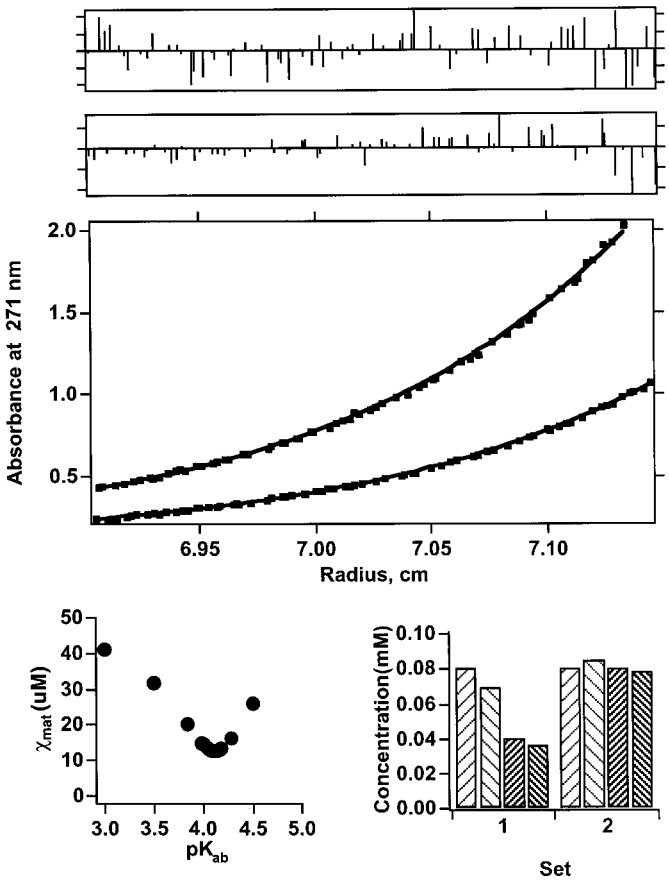


FIG. 4. Results from fitting SP1189/IL-2 AU data with HetFitter. Measurements were taken at equilibrium with detection at 270 nm and a rotor speed of 25K rpm. For SP1189, the following data were input by the user: MW = 453 Da, $\epsilon_a = 1250\ M^{-1}\ cm^{-1}$. For IL-2, user input data included MW = 16,000, $\epsilon_b = 9800\ M^{-1}\ cm^{-1}$. For both data sets, $\nu = 0.753$ ml/g and $\rho = 1.002$ g/ml. Loading concentrations were 80 μM for SP1189 (both data sets) and 40 and 80 μM for IL-2 (sets 1 and 2, respectively). Cell meniscus positions were fixed to the first point used in the data set, and bottom position was determined graphically based on loss of light intensity at the chamber boundary. (Top panels) The fit residuals at the optimum pK_{ab} for sets 1 and 2. (Center panels) The raw data points and optimum pK_{ab} fit lines. (Bottom panels) Left, the RMS deviations of computed concentrations (χ_{mat}) versus assumed pK_{ab}; and right, histograms comparing actual (downward left hatching) and computed (downward right hatching) concentrations for components A (light) and B (dark) for each set obtained with fitting at optimum pK_{ab} value.

units, adjusted extinction coefficients, and loading concentrations we used in HetFitter (Fig. 6). In contrast to the data described for HetFitter (Fig. 3), 2C1SFIT not only provided very unsatisfactory fits to both data sets but also, for the 290-nm data, gave a pK_{ab} value (10.8) well outside the range expected (4–9). The pK_{ab} value for the 330-nm data (5.1) was within this range but the poor quality of the data fit would preclude attaching any significance to this value. Allowing the components to have globally varying association "competencies" (the fcomp parameters in the program) frequently gave nonconverging curve fits or good fits for only a single data set, and fits were highly dependent on parameter starting values. In general, varying these parameters gave unsatisfactory results.

The underlying reason for the inability of the TWOCOMP program to fit this data is that small errors in concentrations excessively bias the curve fit. To demonstrate this more clearly, we used simulated equilibrium sedimentation data where heterodimerization was assumed to occur between monomer A (200 Da) and monomer B (15 kDa) with a pK_{ab} = 5. Three different molar concentrations of components were chosen to simulate situations where equilibrium concentration profiles would be sensitive to values of pK_{ab}. These data sets were then fit globally using both HetFitter and 2C1SFIT. In fitting with HetFitter, pK_{ab} was fixed to values spanning ±2 pK_{ab} units of the simulation value (pK_{ab} = 5). In fitting with 2C1SFIT pK_{ab} was allowed to vary freely, but with material constraints on the two component concentrations. With exactly correct component concentrations both methods returned equally good fits to the data (not shown)

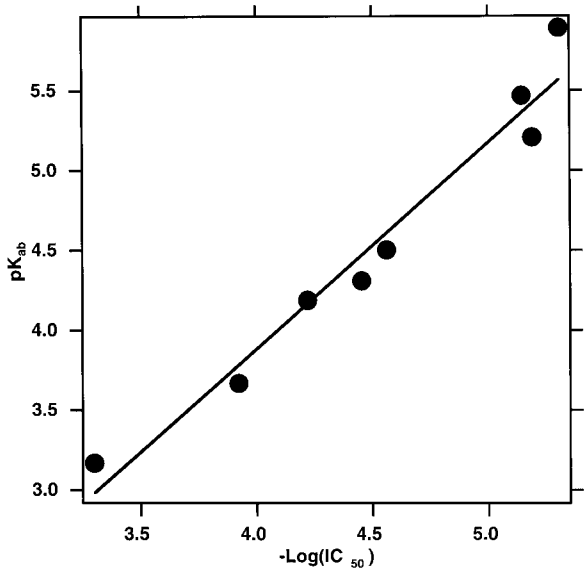


FIG. 5. Correlation plot showing pK_{ab} values calculated from AU data plotted versus $-\log(IC_{50})$ data determined by SPA for a series of compounds (see supplementary material for details). AU measurements were performed on 25, 40, or 80 mM compound; the concentration was chosen based on the absorptivity of the molecule. Typically, three concentrations of protein were added to a constant concentration of compound, at IL-2:compound ratios of 1:4, 1:1, and 1:4. Determinations of K_{ab} were made at rotor speeds of 20K and/or 25K rpm. Inhibition constants (IC_{50}) were measured in SPA experiments where compound concentrations were varied 200,000-fold. Reported values represent an average of at least two experiments and are accurate within twofold of the average. While we have fewer trials for AU data, the accuracy for K_{ab} values in HetFitter seems also to be within twofold of the average. The line represents the best fit to the data and has a slope of 1.29 and a correlation coefficient of 0.98.

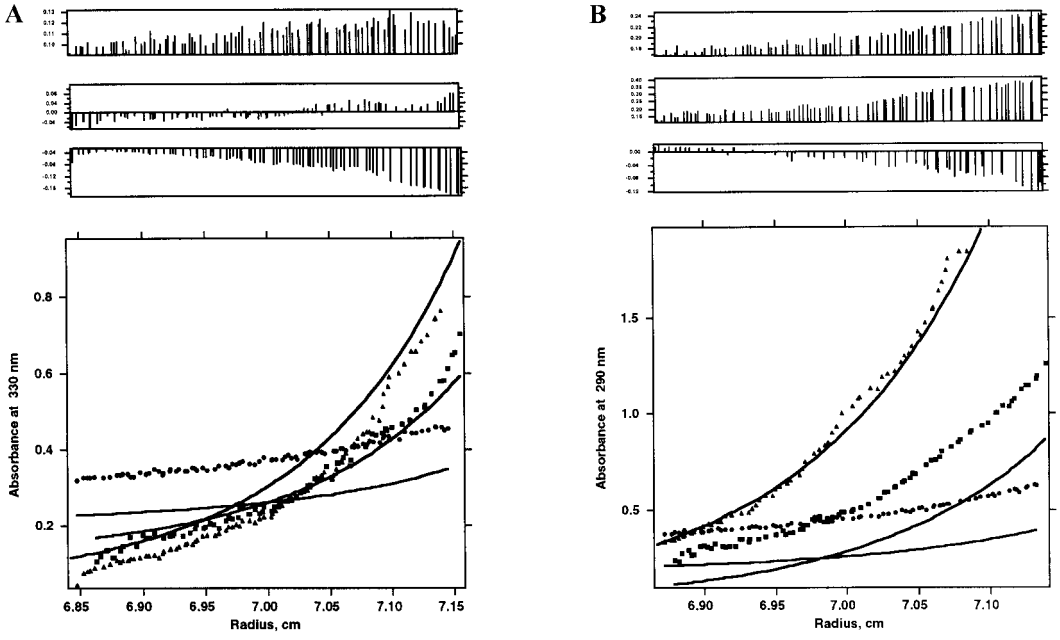


FIG. 6. Results of using "2C1SFIT" (TWOCOMP) to fit data sets of Fig. 3. Parameters used were the same (after required units conversions) as with HetFitter. Only the monomer heteroassociation constant (K11 in 2CSFIT) was allowed to vary for fitting. Values returned (converted to units used in HetFitter) were $pK_{ab} = 4.7$ (330-nm data in A) and $pK_{ab} = 10.8$ (290-nm data in B).

and the correct value for the heterodissociation constant. However, when the concentrations of component A were in error by ±10% (see Fig. 7), 2C1SFIT provided a very unsatisfactory fit to the data whereas HetFitter returns the same quality (actually, exactly the same) fit to the experimental data and also the correct value for pK_{ab}. The χ_{mat} returned for pK_{ab} in HetFitter is higher when the concentrations are in error by 10%; this increased value of χ_{mat} provides a diagnosis for errors in the user-input data. Interestingly, for these simulated data sets, the value for pK_{ab} returned by 2C1SFIT (4.95) was very close to correct. This is because, in this particular, deliberately chosen case, the concentration of the virtually invisible component B could be anything were it not for the material constraint. This constraint forces the equilibrium calculation of 2C1SFIT to be close to correct, but at the expense of the fit quality. This was also the case with the 2C1SFIT fitting of 330-nm data for IL-2/SP2214 binding (Fig. 6A). For the 290-nm data, where both components contribute to the signal, highly erroneous pK_{ab} results were obtained.

Notes on Accuracy and Reliability

We have shown that equilibrium AU can be used to obtain K_{ab} values for small-molecule ligands of IL-2. Since the early observations of Schachman (12), there have been very few examples of such experiments in the literature, and therefore HetFitter could be an important new tool in the drug-discovery laboratory. However, as with any quantitative method employing complex, nonlinear calculations, the reliability of the calculated quantities must be assessed carefully. For example, when binding is so tight that negligible quantities of either monomeric component are present, one can only determine the lower limit for the dissociation constant. In the HetFitter procedure, this instance shows up as a plateau instead of a well-defined minimum in χ_{mat} vs pK_{ab} plot. Because most compound and protein extinction coefficients are less than 20,000 M^{-1} cm^{-1}, measurable (>0.02 A) concentrations of components are limited to the micromolar range so submicromolar binding constants cannot accurately be determined. Development of fluorescence-detected AU (24) would permit measurement of much tighter binders.

It is helpful, where possible, to design experiments by mathematical modeling to define optimum ranges of experimental parameters. The great versatility of UV absorption at user-selectable wavelengths allows component concentrations to be selected for optimum measurement of binding constants. For example, as we have done with the series of compounds reported above, we were able to choose concentration ranges and appropriate wavelengths to observe binding to IL-2 with K_{ab} values ranging from 5 to 500 μM. Of course, the accuracy of these dissociation constants is limited by the accuracy of extinction coefficient and baseline data. We (and Minton (11)) have found that best fitting

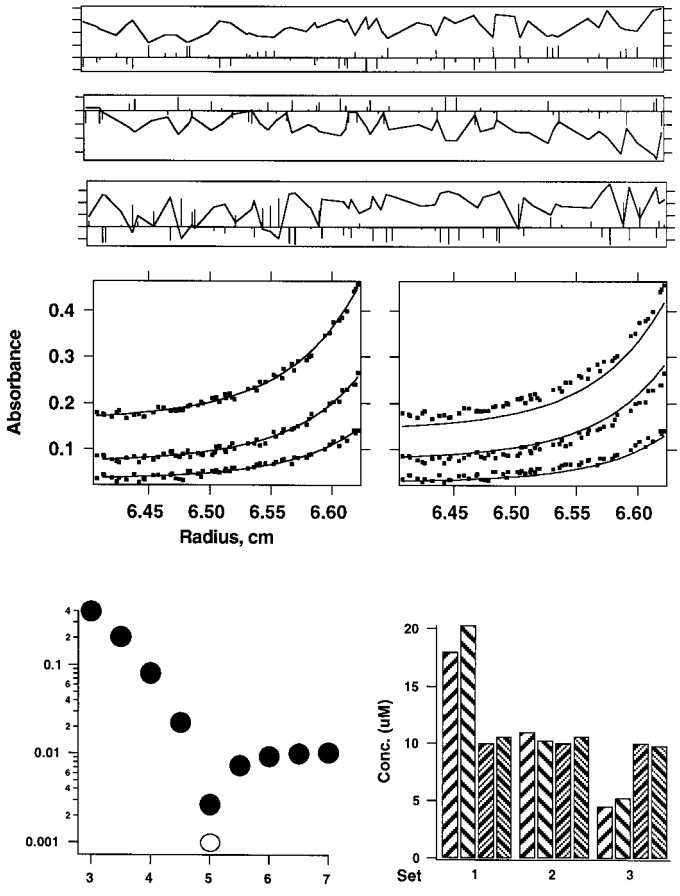


FIG. 7. Results of fitting simulated data sets for MW_A = 200; MW_B = 15,000; ϵ_a = 10,000; ϵ_b = 100; simulation loading concentrations of component A = 20, 10, and 5 μM; program input concentrations = 18, 11, and 4.5 μM. Component B simulation and input concentrations were 100 μM (all sets). Cell meniscus and bottom positions were fixed to the exact values used in the simulations. (Top) Optimum fit residuals for HetFitter (bars) and 2CSFIT (lines; these latter have *y* values reversed from HetFitter's for clarity of presentation); (center) data points and fitted lines for HetFitter (left) and 2CSFIT (right); (bottom left) RMS deviations of computed concentrations (χ_{mat}) versus pK_{ab}; and (bottom right) histograms of actual and computed concentrations for each set obtained with HetFitter fitting at optimum pK_{ab} value.

(lowest χ_{mat} in HetFitter) is obtained when extinction coefficients are derived from XLA absorbance data. As we demonstrate above, acceptable curve fits and pK_{ab} values are not overly sensitive to concentration errors when experimental conditions are chosen judiciously. However, as in other equilibrium centrifuge methods, incorrect baselines can have a significant effect on the accuracy of parameters determined by curve-fitting of concentration gradient data. For experiments using low-molecular-weight compounds that contribute significantly to absorbance at the meniscus, baselines should not be allowed to vary in curve fitting. Instead, we assume them to be zero and routinely run a meniscus depletion experiment with samples containing only protein to check the accuracy of that assumption. The sensitivity of the other parameters to baseline errors can be determined by fitting data for a reasonable range of fixed baseline values. For the data in Fig. 3A, for example, fitting with fixed baselines of $\pm$0.01 absorbance units produced a variation of $\pm$0.13 in minimum pK_{ab} values.

CONCLUSION

HetFitter is a method for analyzing heteromeric dissociation constants. We have demonstrated its use for measuring the binding of small-molecule ligands to proteins; however, the method is general and can be used for analyzing protein–protein and protein–nucleic acid interactions. HetFitter handles mass balance in a separate step from curve fitting, distinguishing this method from previously published approaches. Treating mass balance as a secondary parameter allows for the inevitable difficulties with determining compound concentrations and extinction coefficients accurately. Inaccuracies in these user-input parameters result from errors in measuring small amounts of material as well as from limitations in the wavelength accuracy of the XLA itself. Calculated K_{ab} values are not highly sensitive to small errors in concentration, and the value of χ_{mat} at the K_{ab} provides an additional check of user-input parameters. HetFitter is therefore a robust method for measuring K_{ab}'s for small-molecule/protein interactions.

AU methods have been receiving increased attention from the drug-discovery community, and improvements in both hardware and software continue to expand the range of uses for this technique. AU is able to detect binding of suitably light-absorbing compounds to proteins with K_{ab}'s as low as 1 μM, making it an appropriate method for early stage drug discovery. Centrifugation data has been used qualitatively to rank the affinity of ligands as well as identify non-drug-like mechanisms of inhibition. HetFitter further expands the utility of AU for drug discovery by permitting quantitative assessment of drug/protein interactions.

ACKNOWLEDGMENTS

We thank Jennifer Hyde for SPA measurements and Johan Oslob for synthesis of organic compounds. This work was supported in part by NIH Grant 60610-02 to J.D.L.

REFERENCES

1. Hensley, P., and Smith, L. M. (1997) Analytical biotechnology. *Curr. Opin. Biotechnol.* **8,** 1–5.
2. Minton, A. P. (1990) Quantitative characterization of reversible molecular associations via analytical centrifugation. *Anal. Biochem.* **190,** 1–6.

3. Schachman, H. K. (1959) Ultracentrifugation in Biochemistry, Academic Press, New York.
4. Hansen, J. C., Lebowitz, J., and Demeler, B. (1994) Analytical ultracentrifugation of complex macromolecular systems. *Biochemistry* **33,** 13155–13163.
5. Rivas, G., Tangemann, K., Minton, A. P., and Engel, J. (1996) Binding of fibrinogen to platelet integrin alpha IIb beta 3 in solution as monitored by tracer sedimentation equilibrium. *J. Mol. Recognit.* **9,** 31–38.
6. Rivas, G., Stafford, W., and Minton, A. P. (1999) Characterization of heterologous protein–protein interactions using analytical ultracentrifugation. *Methods* **19,** 194–212.
7. Boehm, H. J., Boehringer, M., Bur, D., Gmuender, H., Huber, W., Klaus, W., Kostrewa, D., Kuehne, H., Luebbers, T., Meunier-Keller, N., and Mueller, F. (2000) Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization—A promising alternative to random screening. *J. Med. Chem.* **43,** 2664–2674.
8. Koepf, E. K., Petrassi, H. M., Ratnaswamy, G., Huff, M. E., Sudol, M., and Kelly, J. W. (1999) Characterization of the structure and function of W → F WW domain variants: Identification of a natively unfolded protein that folds upon ligand binding. *Biochemistry* **38,** 14338–14351.
9. Niederhauser, O., Mangold, M., Schubenel, R., Kusznir, E. A., Schmidt, D., and Hertel, C. (2000) NGF ligand alters NGF signaling via p75(NTR) and trkA. *J. Neurosci. Res.* **61,** 263–272.
10. Lewis, M. S., Schrager, R. I., and Kim, S.-J. (1994) *in* Modern Analytical Ultracentrifugation (Schuster, T. M., and Laue, T. M., Eds.), pp. 94–115, Birkhåuser, Boston, MA.
11. Minton, A. P. (1997) Alternative strategies for the characterization of associations in multicomponent solutions via measurement of sedimentation equilibrium. *Prog. Colloid Polymer Sci.* **107,** 11–19.
12. Schachman, H. K., and Edelstein, S. J. (1966) Ultracentrifuge studies with absorption optics. IV. Molecular weight determinations at the microgram level. *Biochemistry* **5,** 2681–2705.
13. Wills, P. R., Jacobsen, M. P., and Winzor, D. J. (2000) Analysis of sedimentation equilibrium distributions reflecting nonideal macromolecular associations. *Biophys. J.* **79,** 2178–2187.
14. Wilson, E. K., Scrutton, N. S., Colfen, H., Harding, S. E., Jacobsen, M. P., and Winzor, D. J. (1997) An ultracentrifugal approach to quantitative characterization of the molecular assembly of a physiological electron-transfer complex: The interaction of electron-transferring flavoprotein with trimethylamine dehydrogenase. *Eur. J. Biochem.* **243,** 393–399.
15. Winzor, D. J., Jacobsen, M. P., and Wills, P. R. (1998) Direct analysis of sedimentation equilibrium distributions reflecting complex formation between cytochrome c and ovalbumin. *Biochem. Soc. Trans.* **26,** 741–745.
16. Kim, S. J., Tsukiyama, T., Lewis, M. S., and Wu, C. (1994) Interaction of the DNA-binding domain of *Drosophila* heat shock factor with its cognate DNA site: A thermodynamic analysis using analytical ultracentrifugation. *Protein Sci.* **3,** 1040–1051.
17. Tilley, J. W., Chen, L., Fry, D. C., Emerson, S. D., Powers, G. D., Biondi, D., Varnell, T., Trilles, R., Guthrie, R., Mennona, F., Kaplan, G., LeMahieu, R. A., Carson, M., Han, R., Liu, C., and Palermo, R. (1997) Identification of a small molecule inhibitor of the IL-2/IL-2R receptor interaction which binds to IL-2. *J. Am. Chem. Soc.* **119,** 7589–7590.
18. Arkin, M. R., Randal, M., Oslob, J., Wells, J., *et al.*, manuscript in preparation.
19. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) Analysis of data from the analytical ultracentrifuge by nonlinear least-squares techniques. *Biophys. J.* **36,** 575–588.
20. Williams, J. W., van Holde, K. E., Baldwin, R. L., and Fujita, H. (1958) The theory of sedimentation analysis. *Chem. Rev.* **58,** 715–744.
21. Laue, T., Shaw, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) *in* Analytical Ultracentrifugation in Biochemistry and Polymer Science (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.), pp. 90–125, Royal Soc. Chem., Cambridge, UK.
22. Devos, R., Plaetinck, G., Cheroutre, H., Simons, G., Degrave, W., Tavernier, J., Remaut, E., and Fiers, W. (1983) Molecular cloning of human interleukin 2 cDNA and its expression in *E. coli*. *Nucleic Acids Res* **11,** 4307–4323.
23. Landgraf, B. E., Williams, D. P., Murphy, J. R., Smith, K. A., and Ciardelli, T. L. (1991) Conformational perturbation of interleukin-2: A strategy for the design of cytokine analogs. *Proteins* **9,** 207–216.
24. Crepeau, R. H., Conrad, R. H., and Edelstein, S. J. (1976) UV laser scanning and fluorescence monitoring of analytical ultracentrifugation with an on-line computer system. *Biophys. Chem.* **5,** 27–39.