Methods 54 (2011) 39–55

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Review Article

Evaluating the stoichiometry of macromolecular complexes using multisignal sedimentation velocity

Shae B. Padrick, Chad A. Brautigam *

Department of Biochemistry, The University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd., Dallas, TX 75390-8816, United States

ARTICLE INFO

Article history:
Available online 20 January 2011

Keywords:
Analytical ultracentrifugation
Biophysical methods
Stoichiometry
Multisignal sedimentation velocity
Arp2/3 complex
Step by step instructions

ABSTRACT

Gleaning information regarding the molecular physiology of macromolecular complexes requires knowledge of their component stoichiometries. In this work, a relatively new means of analyzing sedimentation velocity (SV) data from the analytical ultracentrifuge is examined in detail. The method depends on collecting concentration profile data simultaneously using multiple signals, like Rayleigh interferometry and UV spectrophotometry. If the cosedimenting components of a complex are spectrally distinguishable, continuous sedimentation-coefficient distributions specific for each component can be calculated to reveal the molar ratio of the complex's components. When combined with the hydrodynamic information available from the SV data, a stoichiometry can be derived. Herein, the spectral properties of sedimenting species are systematically explored to arrive at a predictive test for whether a set of macromolecules can be spectrally resolved in a multisignal SV (MSSV) experiment. Also, a graphical means of experimental design and criteria to judge the success of the spectral discrimination in MSSV are introduced. A detailed example of the analysis of MSSV experiments is offered, and the possibility of deriving equilibrium association constants from MSSV analyses is explored. Finally, successful implementations of MSSV are reviewed.

1. Introduction

Knowledge of the stoichiometry of the components of macromolecular assemblies is fundamental to understanding their physiology. A host of methods to determine this quantity is available to the modern researcher. Among these methodologies are electron microscopy, X-ray crystallography, NMR spectroscopy, calorimetry, and light scattering. Analytical ultracentrifugation (AUC) is naturally applied to this problem because the data report directly on the molar mass (sedimentation equilibrium, SE) or size and shape (sedimentation velocity, SV) of the macromolecular complex.

Five years ago, Schuck and others introduced a new treatment for SV data obtained from multiple signals [1]. These signals can be different wavelengths obtained from the absorbance optical system of the ultracentrifuge and/or interferometric data obtained from the on-board Rayleigh interferometer. Usually, these signals are obtained simultaneously during a single SV experiment. Here, we term this methodology Multi-Signal SV (MSSV). The analysis of MSSV data decomposes the standard $c(s)$ distribution into component distributions called $c_k(s)$, where the k denotes an individual component present in solution. Because individual components can be detected using MSSV, it is possible to analyze the populations of 2–4 cosedimenting components. This method is therefore useful for the analysis of non-interacting species that exhibit little or no hydrodynamic resolution [1]. However, MSSV has found most utility in the analysis of cosedimenting interacting species. Examination of the relative populations of such species, coupled with the complex's hydrodynamic properties, allows the experimenter to determine the complex stoichiometry. This quantity is vital for understanding the functioning of the macromolecular complex, and it is a primary focus of many other analytic methods, e.g. isothermal titration calorimetry.

In this paper, we detail the experimental aspects of the MSSV method. After an introduction to the theoretical basis for the method, we then explore some critical experimental prerequisites that should be met before the experimenter performs an MSSV study. Among these considerations are empirical assessments of sample quality and suitability. Also, we develop pre-analysis processes designed to answer the questions (1) can MSSV work for my macromolecules? and (2) what concentrations of components should I use to maximize the probability of success? We also introduce a binding isotherm that is uniquely available from the MSSV analy-

Abbreviations: Arp2/3, actin related protein 2–actin related protein 3 complex; AUC, analytical ultracentrifugation; LE, Lamm equation; MSSV, multisignal sedimentation velocity; OD, optical density; SE, sedimentation equilibrium; SV, sedimentation velocity analytical ultracentrifugation; VCA, verprolin homology – central region – acidic region; f_r, frictional ratio; r.m.s.d., root-mean-square deviation.

* Corresponding author. Fax: +1 214 645 5383.
E-mail address: chad.brautigam@utsouthwestern.edu (C.A. Brautigam).

1046-2023/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymeth.2011.01.002

sis. A detailed description of the analysis of two interacting proteins is presented. Finally, we discuss the implications of the present work and the significant biological impact that MSSV analysis has made thus far.

2. Methods

2.1. Protein methods

The fusion of glutathione-S-transferase and the VCA domain of WASP (GST-VCA) was constructed and purified as detailed in Ref. [2]. Arp2/3 complex was isolated from bovine thymus as described [2]. The proteins were dialyzed against a buffer comprising 50 mM KCl, 10 mM imidazole, 1 mM EGTA (ethylene glycol-bis(2-aminoethylether)-N,N,N′,N′-tetraacetic acid), and 1 mM $MgCl_2$. The samples were placed in an assembled ultracentrifugation cell that had a charcoal-filled Epon, 1.2-cm-path-length, two-sectored centerpiece sandwiched between sapphire windows. Three samples were placed in an An-60Ti rotor and were centrifuged at 42,000 rpm in an Optima XL-I ultracentrifuge (Beckman-Coulter, Fullerton, CA): GST-VCA alone, Arp2/3 alone, and a mixture of the two proteins. Data were acquired using the absorbance optics (280 nm) and the interference optics.

2.2. Data analysis

All data were analyzed using SEDPHAT version 8.2. Because the refinement of parameters can be dominated by data sets with significantly more data points, a new feature of SEDPHAT was utilized that compensates for this imbalance. This is the "*sqrt(N1/Nx)" checkbox in the Experimental Parameter dialog box. This feature was activated in all analyses presented in this work (except where noted), and a more detailed explanation of its basis and workings is found in [3].

2.3. Simulation of MSSV data for Section 4

Data were simulated in SEDFIT and SEDPHAT. First, a mock data set with the proper time and radial resolution was produced using the generate function in SEDFIT. Default values were used for the analysis in SEDFIT, except the meniscus was set to 6 cm and radial resolutions of 0.003 cm and 0.00072 cm were used for absorbance and IF data, respectively. Except where noted, time resolution was one trace per five minutes. Rotor acceleration was modeled using the default settings. Next, the SV data set was created in SEDPHAT. The mock data set was loaded into SEDPHAT, the meniscus was fixed at 6 cm and the sample bottom at 7.2 cm, and baseline fitting was deactivated. For all steps of the simulation, default values for pathlength, (1.2 cm), $\bar{v}$ (0.73 ml/g), buffer density (0.998230 g/ml), buffer viscosity (0.010020 P), and temperature (20 °C) were used. The A + B ↔ AB model was selected. Parameters were entered as follows: Component A: $[A]_{tot}$ = 1.9 μM, molar mass = 200,000 g/mol, s_A = 8.5 S, ε^A_{ABS} = 140,845 M^{-1} cm^{-1} or ε^A_{IF} = 550,000 fringes·M^{-1} cm^{-1}, depending on whether absorbance or interference data was being modeled. Component B: $[B]_{tot}$ = 4.2 μM, molar mass = 10,000 g/mol, s_B = 1.2 S, ε^B_{IF} = 27,500 fringes·M^{-1} cm^{-1} when interference data was being modeled, and ε^B_{ABS} was set to 25000, 20000, 16000, 14000, 12000, 11000, 10000, 9000, 8000, 7700, 7500, or 7300 M^{-1} cm^{-1}, with the value of ε^B_{ABS} varied to generate the range of D_{norm} (see Eq. (9)) seen in Fig. 2A. The complex AB parameters: s_{AB} = 9.2 S, $\log(K_A)$ = 9 and $\log(k_{off})$ = −3. Conservation of mass was switched on. Synthetic noise was added to the simulated SV data and the data was saved. This was performed for each signal used. Finally, the multiple-signal data sets were analyzed a new SEDPHAT session, following steps analogous to steps 20–26 of the Supplemental Protocol (see Section 5.1). Absorbance signals were given a noise value of 0.0049 to account for this different number of data points in the IF and absorbance data sets, simulated IF data noise level was left at the default, 0.01 value. The multisignal sedimentation velocity model was chosen, and Marquardt–Levenberg fitting algorithm was chosen. The global analysis was set up with the known signal increments (from the design of different D_{norm}'s), so there was no need to optimize these. Two regions of s-space were defined, one for free excess components, the other for the complex. In each, a spectrum (i.e. a $c_k(s)$ distribution) for each component was calculated. The frictional ratio (f_r) was optimized in each. Initially, a single global run was performed to optimize linear parameters, then a single round of fitting was performed, optimizing only the frictional ratios. Unlike in steps 20–26 of the example, the menisci, TI and RI noise were not fit. Instead the cell bottom and menisci were input from the known parameters from the modeling step (7.2 and 6 cm, respectively). Concentrations of different species sedimenting at different rates were quantified by integrating a range around relevant peaks.

Three-component/three-signal systems were models similarly. For simplicity, the simulated SV data for the component system was modeled using the A + B ↔ AB system described above, with $\log(K_A)$ = 9 and $\log(k_{off})$ = −3. In this case, "B" in the SV data generation is a tight complex of components B and C (as analyzed in the second step of Section 4.1). Sedimentation coefficients were chosen to be: System (#1): s_A = 8.5 S, s_B = 2, s_{AB} = 9.2; System (#2): s_A = 3.5 S, s_B = 4.0 S, s_{AB} = 6.0 S; System (#3): s_A = 3.5 S, s_B = 4.0 S, s_{AB} = 6.5 S; System (#4): s_A = 3.5 S, s_B = 4.0 S, s_{AB} = 6.0 S; System (#5): s_A = 8.5 S, s_B = 3.5 S, s_{AB} = 9.5 S. Relevant masses and extinction coefficients are given in Table 1, calculated total concentrations (by summing the quantities in the two peaks) are given in Table 2. MSSV analysis was performed as described for the two-signal, two-component system, but using three wavelengths.

3. Theory

3.1. The c(s) distribution

This work is primarily concerned with the modeling of SV data. The concentration profiles obtained over time in an SV experiment may be directly modeled as the integral (or discretized sum) of solutions to the Lamm Equation (LE) scaled by a continuous distribution called $c(s)$ [4,5]. If $a(r,t)$ represents the data acquired from an SV experiment, then

$$a(r,t) \cong \int_{s_{min}}^{s_{max}} c(s)\chi(s, D(s), r, t)ds, \quad (1)$$

where s is the sedimentation coefficient and χ (s, $D(s)$, r, t) is a LE solution that is dependent on $D(s)$, the corresponding diffusion coefficient, r, radius from the center of rotation, and t, the time from the beginning of the experiment. Several important features of this type of analysis must be pointed out. First, the analysis directly fits the SV data, i.e. no modifications to the data like pairwise subtractions of scans are needed. Aiding this is the ability to accurately model the noise structure of the concentration profiles [6]. Another notable feature of this analysis is that the LE solutions are for ideal, non-interacting species. Despite this fact, the diffusional deconvolution afforded by the $c(s)$ analysis allows the accurate description of the mass-transport properties [7,8] and apparent diffusion coefficients for reaction boundaries [9] for interacting systems. Usually, a single f_r is assumed over the entire distribution. However, the distribution can be divided into segments, with each segment having its own f_r. This approach is useful in cases where two or more boundaries are easily identified in the data (as in Section 5.1.3, below) [10]. In addition, many frictional ratios can be considered in an