



Keywords: small-angle neutron scattering; contrast variation; *SASSIE-web* platform; macromolecular complex modeling; simulation; software.

Software for small-angle neutron scattering contrast variation experiment planning and data analysis

Susan Krueger,^{a,b} Kathryn L. Sarachan,^c Emre Brookes^d and Joseph E. Curtis^{b*}

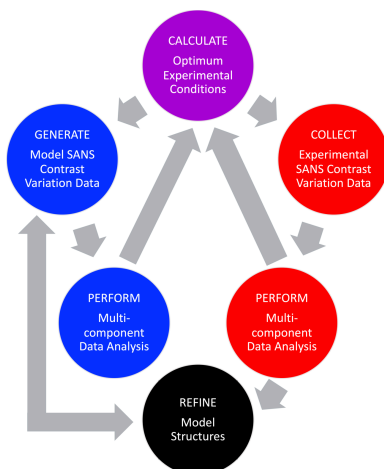
^aDepartment of Materials Science and Engineering, University of Maryland, College Park, Maryland, USA, ^bCenter for Neutron Research, National Institute of Standards and Technology, Gaithersburg, Maryland, USA, ^cDepartment of Chemistry, Wilson College, Chambersburg, Pennsylvania, USA, and ^dDepartment of Chemistry, University of Montana, Missoula, Montana, USA. *Correspondence e-mail: joseph.curtis@nist.gov

Contrast variation (CV) small-angle neutron scattering (SANS) enables the determination of structural parameters from individual components in multi-component biological complexes in solution. Researchers can selectively highlight different parts of a multi-component system by varying the contrast between the components, typically through adjusting the deuteration level of solvents or specific molecular constituents. While this approach provides unique insights into component shapes, spatial arrangements and interactions, SANS CV experiments require substantial resources, specialized facilities and, often, a steep learning curve. To address these challenges and improve the likelihood of a successful experiment, a new contrast software suite has been developed within the *SASSIE-web* framework, allowing researchers to plan and simulate experiments *in silico*, integrate structural modeling, and streamline both experiment design and CV data analysis into a unified platform.

1. Introduction

Small-angle neutron scattering (SANS) contrast variation (CV) experiments have long been a powerful technique for studying biological complexes. They enable detailed structural information to be obtained about multi-component systems in solution. By varying the neutron scattering length density of different components, researchers can isolate the scattering signals of individual components, making this method particularly useful for studying complex systems where individual structures are otherwise difficult to discern.

Comprehensive discussions on SANS CV techniques are available in recent reviews and tutorials (Zaccai *et al.*, 2016; Mahieu & Gabel, 2018; Krueger, 2017, 2022a,b; Trehwella, 2022; Whitten & Jeffries, 2023). In practice, however, implementing SANS CV remains demanding, requiring a substantial investment of time and resources, careful experimental design, and access to specialized facilities for sample preparation and data collection. To help address these challenges, the *SASSIE-web* platform (Curtis *et al.*, 2012; Perkins *et al.*, 2016) already provides modules for planning SANS CV experiments, constructing structural models and generating structural ensembles from experimental data – capabilities that have made it a valuable resource for the biological small-angle scattering community. Building on this foundation, the newly developed *Contrast* software suite introduces dedicated modules explicitly designed for SANS CV experiments on multi-component biological complexes in solution. These modules allow researchers to simulate experiments *in silico* before beam time, analyze CV data within the same environment, and integrate structure modeling for both individual



and ensemble structures, creating a unified and robust platform for planning, execution and interpretation of SANS CV experiments.

2. Contrast variation experiment workflow

The CV experiment-planning and data-analysis workflow is shown in Fig. 1. Ideally, a user would begin the workflow by planning the SANS CV experiment, which involves performing a number of preparatory calculations. Optimal experimental conditions, such as sample concentration and contrasts (the fraction of D₂O versus H₂O in the solvent), can be determined using the *SASSIE-web* Contrast Calculator module (Sarachan *et al.*, 2013). If all-atom starting structure models are available, model SANS data can be calculated at the chosen solvent conditions using the SasCalc module. The new Multi-Component Analysis module can then be used with the model SANS curves to perform the CV analysis *in silico* before the actual experiment. This provides information about the radius of gyration, R_g , values for each component and the distance between their centers of mass. These parameters can also be calculated directly from the model structure using R_g Center of Mass Distance Calculator. Good agreement provides confidence that the contrast conditions have been chosen well, increasing the likelihood of a successful SANS CV experiment.

After the experiment has been planned and SANS data have been collected, Multi-Component Analysis can be used with the experimental data and the results compared with those from the model SANS data. The model structures can

then be refined so that the model SANS curves better represent the experimental curves.

3. Contrast software suite

The *Contrast* software suite is written in Python 3, uses the *NumPy* (<https://numpy.org/>) and *SciPy* (<https://scipy.org/>) libraries, and features an enhanced web graphical user interface (GUI), built using the *GenApp* framework (Brookes *et al.*, 2015), that is accessible from a modern web browser such as *Firefox*, *Safari* or *Chrome*. The software is available for general users at <https://sassie-web.chem.utk.edu/sassie2/> as part of the main *SASSIE-web* application. To ensure long-term maintainability and ease of collaboration, the software is hosted on GitHub, <https://github.com/madscatt/sassie>, allowing for open-source development. Furthermore, it is well documented with thorough explanatory comments embedded in the code, enabling future developers to quickly understand the structure and functionality of the program and ensuring its ongoing development and improvement.

Users create a persistent account on the server and can log in at any time to obtain access to their filesystem. Each screen has its documentation page, which is opened via the DOCS tab on the right side of the screen. The complete documentation can also be accessed from the login screen even if the user is not logged in. A FEEDBACK tab on the right side of each screen allows users to provide feedback and report bugs. Plots are rendered using *Plotly* (Plotly Technologies, 2015), and a link to *Plotly Chart Studio* is included to enable the user to modify and save the output plots. The data used to make the plots are also included in an output file, allowing users to plot the results with their choice of plotting tools.

Once logged in, the user can access the *Contrast* software suite menu and choose one of the *Contrast* modules: (1) Contrast Calculator (Sarachan *et al.*, 2013), (2) Multi-Component Analysis and (3) R_g Center of Mass Distance Calculator. Together, the three modules provide a collection of comprehensive SANS CV experiment-planning and analysis tools. Contrast Calculator calculates the neutron and X-ray scattering length density, contrast and forward scattering, $I(0)$, values for a single molecule or a multi-component complex consisting of proteins and/or nucleic acids, as well as other compounds such as polymers, lipids and carbohydrates. Multi-Component Analysis builds upon the web-based *MULCh* (Whitten *et al.*, 2008) software. Like *MULCh*, Multi-Component Analysis contains sub-modules that focus on the specialized analyses of SANS CV data from two-component complexes. Among several differences, one notable distinction is that Multi-Component Analysis provides both written and graphical outputs, allowing the user to quickly assess the validity of the result. R_g Center of Mass Distance Calculator allows direct comparison of some of the Multi-Component Analysis output parameters with those calculated from an all-atom model structure. The modular design of the *Contrast* software suite integrates smoothly with existing *SASSIE-web* modules, offering a streamlined and user-friendly experience that allows users to enter and exit at various points of the CV

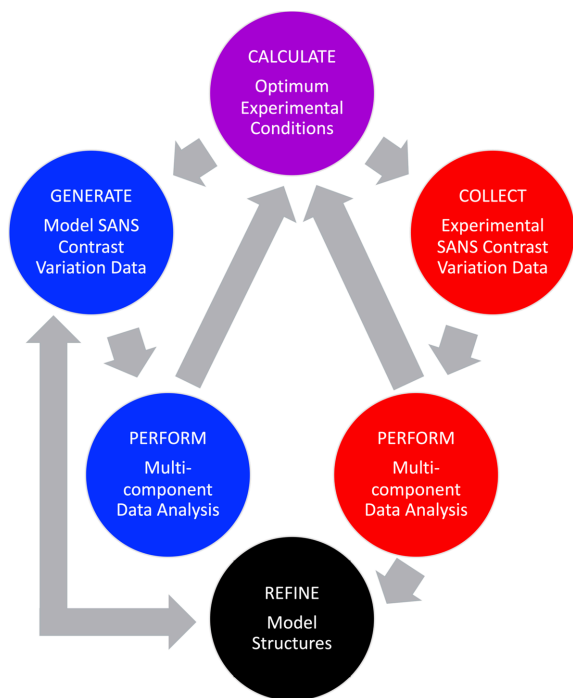


Figure 1
SANS CV experiment workflow. The left side shows the steps for planning a SANS CV *in silico* experiment using model SANS data and the right side shows the corresponding steps using experimental data.

workflow depending on where they are in the process. In particular, the Contrast Calculator module (Sarachan *et al.*, 2013) now writes a general output file containing molecular parameters that can be input into Multi-Component Analysis, eliminating the need to enter them manually.

3.1. Contrast Calculator

Contrast Calculator, now one of the *Contrast* modules, has been described previously (Sarachan *et al.*, 2013). The Contrast Calculator screen is shown in Fig. 2. The icons and labels on the upper right appear on every *SASSIE-web* page and allow the user to keep track of projects, files and jobs as well as customize the work environment. The *Contrast* icon is visible on the upper left of the screen, and the three *Contrast* modules are listed horizontally across the page. Contrast Calculator appears in bold text since it is the selected module. The three vertical lines to the left of the *Contrast* icon allow navigation back to the main *SASSIE-web* screen. In contrast to the previous version of Contrast Calculator, the inputs for the molecule and solvent components are now in a more esthetically pleasing table format. The output now includes estimates of incoherent scattering as a function of the deuteration level of the solvent. This is especially useful for users who are new to SANS CV and may not realize how incoherent scattering affects the measured data.

3.2. Multi-Component Analysis

Multi-Component Analysis is the module that analyzes CV SANS data. It can be used on model data, before an experi-

ment, or on experimental CV data. Dilute solutions are assumed. It is also assumed that the fundamental analyses needed for small-angle scattering data validation (Trehwella *et al.*, 2017; Trehwella, 2023), such as obtaining the radius of gyration, R_g , and the intensity at zero angle, $I(0)$, from Guinier & Fournet (1955) and/or distance distribution function analysis (Glatter, 1977; Moore, 1982; Semenyuk & Svergun, 1991; Svergun, 1992), have been previously completed.

The Multi-Component Analysis module consists of four sub-modules: (1) Match Point, (2) Stuhmann and Parallel Axis Theorem, (3) Decomposition, and (4) Stoichiometry. In each case, the main calculation involves analysis of SANS data obtained at multiple contrasts. The Match Point sub-module can be applied to SANS CV data from complexes with any number of components, whereas the other three sub-modules are applicable to SANS CV data from two-component complexes. Users choose the desired sub-module from the menu on the Multi-Component Analysis screen. The inputs are dependent on the chosen sub-module. Only one analysis can be chosen at a time. Not all the analyses apply to a given set of SANS CV data. It is the responsibility of the user to select the analyses that are applicable to their specific SANS CV dataset.

3.2.1. Match Point Analysis

Match Point Analysis calculates the contrast match point of a multi-component complex from experimental SANS CV data. It is assumed that $I(0)$ has been determined at each

The screenshot shows the SASSIE-web Contrast Calculator interface. At the top, there's a header with the SASSIE-web logo and navigation links: Contrast Calculator (selected), Multi-Component Analysis, and Rg Center of Mass Distance Calculator. On the right, there are links for Logoff, Help on, and a user profile icon. The main form is titled 'Contrast Calculator' and includes several input sections. On the left, there are labels for 'run name', 'output filename prefix (no extension)', 'number of protein, DNA or RNA input files', 'number of additional components', 'Optional Input', 'total solute concentration (mg/mL)', 'solvent % D2O step', 'fraction exchangeable H (protein)', 'fraction exchangeable H (nucleic acid)', and 'number of non-water solvent components'. On the right, there are corresponding input fields: 'run_0' for run name, 'contrast_output' for output filename, '1' for number of input files, '0' for additional components, and a table for optional inputs. The table has columns for 'chemical formula', 'concentration (M)', and 'mass density (g/mL)'. The first row shows 'NaCl', '0.15', and '2.16'. There are also buttons for 'Submit' and 'Reset'. On the far right, there is a vertical menu with 'DOCS' and 'HELP' tabs. At the bottom, a footer line states: 'Current funding provided by NSF grants (CHE-1265821, OAC-1912444 & OAC-1739549) and NIST'.

Figure 2

The Contrast Calculator screen contains the inputs for the Contrast Calculator module. Clicking the DOCS tab on the right opens the Contrast Calculator documentation page, which contains links to the main *SASSIE* and *Contrast* suite documentation pages.

contrast (fraction D₂O in the solvent) and that the concentration of the particles in solution has been measured at each contrast. The result can be directly compared with the complex match point that was calculated by the Contrast Calculator module. Agreement between the two results helps to validate the experimental data.

As shown in Fig. 3, users input $I(0)$ (cm⁻¹), error on $I(0)$ (cm⁻¹), concentration (mg ml⁻¹) and error on concentration (mg ml⁻¹) as a function of fraction of D₂O in the solvent, f_{D_2O} . The f_{D_2O} corresponding to the contrast match point of the multi-component complex and the error on the contrast match point are returned.

The match point analysis is based on the equation for $I(0)$,

$$I(0) = n(\Delta\rho V)^2, \quad (1)$$

where $I(0)$ is the scattering intensity at $q = 0$, $\Delta\rho$ is the contrast of the complex and V is the volume of the complex. The number density of the complex is given by the equation

$$n = c \frac{N_A}{M_w}, \quad (2)$$

where c is the concentration of particles in the solution, N_A is Avogadro's number and M_w is the molecular weight of the complex. Since $\Delta\rho$ is proportional to the fraction of D₂O in the solvent, $I(0) \propto cf_{D_2O}^2$.

To determine the match point of the complex, $I(0)$ is divided by c to normalize for differences in concentration at each contrast. To make the relationship between $I(0)/c$ and f_{D_2O} linear, $\sqrt{I(0)/c}$ versus f_{D_2O} is solved, where the sign of $\sqrt{I(0)/c}$

is changed for the f_{D_2O} values that are less than the match point where $I(0) \approx 0$. This requires an initial guess for the match point, which is determined by performing a second-order polynomial fit to $I(0)/c$ versus f_{D_2O} and then calculating the f_{D_2O} value where the derivative of the polynomial function is zero. A weighted linear fit is performed to $\sqrt{I(0)/c}$ versus f_{D_2O} to obtain the match point from the slope and intercept of the fitted curve using the relation match point = -intercept/slope. The errors on the slope and intercept are obtained from the covariance matrix from the linear fit:

$$\sigma_i = \sqrt{\text{COV}_{ii}}, \quad (3)$$

where COV_{ii} is the diagonal element of the covariance matrix that corresponds to the i th variable.

3.2.2. Stuhmann and Parallel Axis Theorem Analysis

Stuhmann and Parallel Axis Theorem Analysis performs both Stuhmann (Ibel & Stuhmann, 1975) and parallel axis theorem (Engelman & Moore, 1975) analysis to obtain R_g of the individual components in a two-component complex and the distance between their centers of mass. The results from the two methods should be similar and can be compared. As shown in Fig. 4, users supply the partial specific volume, $\bar{v}$ (cm³ g⁻¹), and M_w (kDa) for each of the two components, as well as the experimental R_g (Å) and error on R_g (Å) as a function of fraction D₂O in the solvent, f_{D_2O} . The contrast values, $\Delta\rho$ (10¹⁰ cm⁻²), are also needed for each component as a function of f_{D_2O} . These $\Delta\rho$ values can be obtained from

SASSIE-web

Contrast Calculator **Multi-Component Analysis** Rg Center of Mass Distance Calculator

run name: run_0

output file name: general_output_file.txt

choose utility type: match point analysis

Match Point Analysis Input

number of contrast points: 4

fraction D2O	total concentration	total concentration error	I(0)	I(0) error
0.0	7.7	0.4	0.85	0.01
0.2	7.7	0.4	0.534	0.044
0.85	7.7	0.4	0.013	0.003
1.0	7.7	0.4	0.095	0.002

Advanced Input

Check Box for Advanced Input: ☐

Submit

Reset

progress: _____

percent done: _____

plotly plot: _____

Current funding provided by NSF grants (CHE-1265821, OAC-1912444 & OAC-1739549) and NIST

Figure 3

The Multi-Component Analysis screen showing the inputs for Match Point Analysis, the default sub-module. Clicking the DOCS tab on the right opens the Multi-Component Analysis documentation page, which contains links to the documentation pages of the four sub-modules.

the chemical composition of the complex using the *SASSIE* Contrast Calculator module (Sarachan *et al.*, 2013). The R_g (Å) of each component, the errors on R_g (Å), the distance between their centers of mass (Å) and the error on that distance (Å) from both the Stuhmann and parallel axis theorem analyses are returned.

The Stuhmann analysis is performed by simultaneously fitting R_g^2 versus $1/\Delta\rho$ at each contrast to the Stuhmann equation:

$$R_g^2 = R_m^2 + \frac{\alpha}{\Delta\rho} - \frac{\beta}{(\Delta\rho)^2}, \quad (4)$$

where R_m is the radius of gyration of the complex at infinite contrast, *i.e.* the radius of gyration of an equivalent homogeneous particle. The sign of α indicates whether the component with the higher scattering length density is located toward the inside or the outside of the complex. If $\beta = 0$, the centers of mass of the two components are concentric. R_g is the measured radius of gyration at each contrast (f_{D_2O}) and

$\Delta\rho$ is the calculated contrast of the complex under the same solvent conditions. The term $\Delta\rho$ is calculated from the contrasts of the individual components $\Delta\rho_1$ and $\Delta\rho_2$:

$$\Delta\rho = f_1\Delta\rho_1 + f_2\Delta\rho_2, \quad (5)$$

where $f_1 = V_1/(V_1 + V_2)$ and $f_2 = 1 - f_1$.

The volumes of the components are calculated from their molecular weights and partial specific volumes:

$$V = \bar{v} \frac{M_w}{N_A}. \quad (6)$$

Errors on α , β and R_m^2 are obtained from the covariance matrix from the fit [equation (3)].

The radii of gyration for the two components, R_1 and R_2 , and the distance between their centers of mass, D , are obtained from α , β and R_m^2 by simultaneously solving (Moore, 1982; Olah *et al.*, 1994)

$$R_m^2 = f_1R_1^2 + f_2R_2^2 + f_1f_2D^2, \quad (7)$$

SASSIE-web

Contrast Calculator **Multi-Component Analysis** Rg Center of Mass Distance Calculator

run name: run_0
output file name: general_output_file.txt
choose utility type: stuhmann and parallel axis theorem analysis

Stuhmann and Parallel Axis Theorem Analysis Input

number of contrast points: 4

fraction D2O	radius of gyration	radius of gyration error
0.0	25.45	0.07
0.2	24.95	0.09
0.85	28.0	3.0
1.0	31.34	0.4

manual data input: 2

name	partial specific volume	molecular weight
component_name_1	0.73	14.3
component_name_2	0.73	44.2

enter delta rho values for each combination

fraction D ₂ O	component_name_1	component_name_2
0.0	2.551	5.104
0.2	1.383	3.928
0.85	-2.415	0.109
1.0	-3.292	-0.773

Advanced Input

Check Box for Advanced Input: ☐

Submit Reset

progress:
percent done:
plotly plot:

Current funding provided by NSF grants (CHE-1265821, OAC-1912444 & OAC-1739549) and NIST

Figure 4

The Multi-Component Analysis screen showing the inputs for the Stuhmann and Parallel Axis Theorem Analysis sub-module. Clicking the DOCS tab on the right opens the Multi-Component Analysis documentation page, which contains links to the documentation pages of the four sub-modules.

$$\alpha = (\Delta\rho_1 - \Delta\rho_2)f_1f_2[R_1^2 - R_2^2 + (f_2^2 - f_1^2)D^2] \quad (8)$$

and

$$\beta = (\Delta\rho_1 - \Delta\rho_2)^2 f_1^2 f_2^2 D^2. \quad (9)$$

Errors on R_1 , R_2 and D are obtained using a Monte Carlo error analysis to propagate the errors on α , β and R_m^2 , which are determined from the fit to the Stuhmann equation. The reported R_g and center-of-mass distance values are average values based on multiple calculations in which α , β and R_m^2 are varied randomly within the range defined by their input errors. The reported errors are standard errors of that mean.

R_1 , R_2 and D are also found by simultaneously fitting the experimental R_g^2 values to the parallel axis theorem at each contrast:

$$R_g^2 = \frac{\Delta\rho_1 V_1}{\Delta\rho V} R_1^2 + \frac{\Delta\rho_2 V_2}{\Delta\rho V} R_2^2 + \frac{\Delta\rho_1 \Delta\rho_2 V_1 V_2}{(\Delta\rho V)^2} D^2, \quad (10)$$

where $\Delta\rho V = \Delta\rho_1 V_1 + \Delta\rho_2 V_2$. In this case, errors on R_1 , R_2 and D are obtained from the covariance matrix from the fit [equation (3)].

3.2.3. Decomposition Analysis

Decomposition Analysis calculates the composite scattering intensities I_{11} , I_{12} and I_{22} for a two-component complex from a set of SANS CV data (Whitten *et al.*, 2008). I_{11} and I_{22} are the

SASSIE-web Logoff Help on

Contrast Calculator **Multi-Component Analysis** Rg Center of Mass Distance Calculator

Multi-Component Analysis

run name: run_0
 output file name: general_output_file.txt
 choose utility type: decomposition analysis

Decomposition Analysis Input

number of contrast points: 4

fraction D2O	data file name	total concentration	total concentration error	starting data point for Guinier analysis	initial number of data points for Guinier analysis
0.0	Browse local files[1] or Browse server	7.7	0.4	1	6
0.2	Browse local files[2] or Browse server	7.7	0.4	1	6
0.85	Browse local files[3] or Browse server	7.7	0.4	1	6
1.0	Browse local files[4] or Browse server	7.7	0.4	1	6

q Rg limit for Guinier analysis: 1.3
 refine scale factor: true
 select option: manual data input
 number of components: 2

name	partial specific volume	molecular weight
component_name_1	0.73	14.3
component_name_2	0.73	44.2

enter delta rho values for each combination

fraction D2O	component_name_1	component_name_2
0.0	2.551	5.104
0.2	1.383	3.928
0.85	-2.415	0.109
1.0	-3.292	-0.773

Advanced Input

Check Box for Advanced Input: ☐
 Submit Reset

progress:
 percent done:
 plotly plot:

Current funding provided by NSF grants (CHE-1265821, OAC-1912444 & OAC-1739549) and NIST

Figure 5

The Multi-Component Analysis screen showing the inputs for the Decomposition Analysis sub-module. Clicking the DOCS tab on the right opens the Multi-Component Analysis documentation page, which contains links to the documentation pages of the four sub-modules.

scattering intensities of components 1 and 2, respectively, and I_{12} is the cross-term, which gives an idea of the spatial relationship between the two components. As shown in Fig. 5, users input the partial specific volume, $\bar{v}$ ($\text{cm}^3 \text{g}^{-1}$), and M_w (kDa) for each of the two components as well as the experimental data file name, concentration (mg ml^{-1}) and error on the concentration (mg ml^{-1}) as a function of fraction D_2O in the solvent, $f_{\text{D}_2\text{O}}$. The contrast values, $\Delta\rho$ (10^{10}cm^{-2}), are also needed for each component as a function of $f_{\text{D}_2\text{O}}$. The scattering intensities of each component (cm^{-1}) and the cross-term intensity (cm^{-1}) are returned.

First, the SANS CV data are read to obtain q ($\text{\AA}^{-1}$) (the scattering vector magnitude), $I(q)$ (cm^{-1}) and the error on $I(q)$ (cm^{-1}) at each contrast ($f_{\text{D}_2\text{O}}$). If errors are not detected when the data are read, as in the case of model data, random Gaussian errors are added. Second, the data are rescaled on the basis of the input concentrations to account for differences in concentration as a function of contrast.

A Guinier analysis (Guinier & Fournet, 1955) is then performed to obtain R_g and $I(0)$ as a function of contrast. A Guinier fit to SANS CV data in the range $qR_{g\text{min}}$ to $qR_{g\text{max}}$ is carried out, where $qR_{g\text{min}}$ is defined by the first data point and $qR_{g\text{max}}$ is based on the qR_g limit, both as specified by the user. A linear fit is performed to the Guinier equation:

$$\ln[I(q)] = \ln[I(0)] - \frac{q^2 R_g^2}{3}. \quad (11)$$

If the user desires, the $I(0)$ values are used to refine the scale factors at each contrast to account for possible concentration inaccuracies. In that case, scale factors are calculated using equations (1) and (2). In this case, $\Delta\rho V = \Delta\rho_1 V_1 + \Delta\rho_2 V_2$. Taking the ratio of the concentrations between two different contrasts, i and j ,

$$\frac{c_i}{c_j} = \frac{I(0)_i (\Delta\rho V)_j^2}{I(0)_j (\Delta\rho V)_i^2}. \quad (12)$$

The scale factors are calculated at each contrast with respect to the first dataset in the scattering data array.

Finally, the composite scattering intensities are found using multiple linear regression at each q value to solve the set of simultaneous equations at each contrast:

$$I_{\text{exp}} = \Delta\rho_1^2 I_{11} + \Delta\rho_1 \Delta\rho_2 I_{12} + \Delta\rho_2^2 I_{22}, \quad (13)$$

where I_{exp} represents the rescaled measured scattering data, and $\Delta\rho_1$ and $\Delta\rho_2$ are the contrasts of components 1 and 2, respectively. Errors on the composite scattering intensities are determined from the covariance matrix at each q value [equation (3)].

Figure 6

The Multi-Component Analysis screen showing the inputs for the Stoichiometry Analysis sub-module. Clicking the DOCS tab on the right opens the Multi-Component Analysis documentation page, which contains links to the documentation pages of the four sub-modules.

Figure 7

The R_g Center of Mass Distance Calculator screen contains inputs for the reference PDB file and a trajectory (PDB or DCD) file. Clicking the DOCS tab on the right opens the documentation page for R_g Center of Mass Distance Calculator, which contains a link to the Basis Selection Syntax documentation page that outlines the syntax required to select the different regions of the complex and provides examples.

3.2.4. Stoichiometry Analysis

Stoichiometry Analysis calculates the molecular weights of the components in a complex containing multiple copies of two components. It assumes that a CV series of measurements has been performed and $I(0)$ has been determined at each contrast. As shown in Fig. 6, users input the partial specific volume, $\bar{v}$ ($\text{cm}^3 \text{g}^{-1}$), of the two components, as well as the $I(0)$ (cm^{-1}), error on $I(0)$ (cm^{-1}), concentration (mg ml^{-1}) and error on concentration (mg ml^{-1}) as a function of contrast ($f_{\text{D}_2\text{O}}$). The molecular weights (kDa) of the two components in the complex and their weight and volume fractions are returned. This allows calculation of the number of copies of each component in the complex on the basis of the M_w values for a single copy of each component.

To solve for the molecular weights of the two components, $I(0)$ is written in terms of equations (1), (2) and (6). Subtracting $I(0)$ from both sides and defining the molecular weight for two components as $M_w = M_1 + M_2$, the molecular weights are obtained by solving the set of simultaneous equations at each contrast:

$$x_0 M_1^2 + x_1 M_1 M_2 + x_2 M_2^2 - x_3 M_1 - x_3 M_2 = 0, \quad (14)$$

where

$$x_0 = \Delta \rho_1^2 \bar{v}_1^2, \quad (15)$$

$$x_1 = 2 \Delta \rho_1 \Delta \rho_2 \bar{v}_1 \bar{v}_2, \quad (16)$$

$$x_2 = \Delta \rho_2^2 \bar{v}_2^2 \quad (17)$$

and

$$x_3 = \frac{I(0) N_A}{c}. \quad (18)$$

The reported errors on the molecular weights are obtained using a Monte Carlo error analysis to propagate the errors on $I(0)$ and the concentration. The reported molecular weights are average values based on multiple calculations in which $I(0)$ and concentration are varied randomly within the range defined by their input errors. The reported errors are standard errors of the mean.

3.3. R_g Center of Mass Distance Calculator

R_g Center of Mass Distance Calculator is the module used to calculate, from a Protein Data Bank (PDB) or trajectory (DCD) file, the contrast-independent R_g of each component in a two-component complex and the distance between their centers of mass. The results can be directly compared with those obtained from the Stuhmann and parallel axis theorem analyses, which also return contrast-independent values of the same parameters from experimental SANS CV data. The R_g Center of Mass Distance Calculator screen is shown in Fig. 7. Inputs are the reference PDB file name, trajectory file (PDB or DCD) name, number of components, component names and the basis string that defines the selection basis for each component. The contrast-independent R_g ($\text{\AA}$) values of the two components and the distance between their centers of mass ($\text{\AA}$) are returned.

4. Tutorials

The *SASSIE-web* documentation is a living resource that is meant to grow and evolve on the basis of user feedback and needs. It includes a Basic Usage tutorial for navigating *SASSIE-web* using the icons and labels at the top of every web page. Examples showing how to keep track of projects, files and jobs are included, along with instructions on how to

customize the work environment. The *SASSIE-web* documentation can be accessed from the login screen without logging in. Once logged in, users can access the documentation from any screen via the DOCS tab on the right.

The documentation for each *Contrast* module includes a tutorial with one or more examples of how to run the module. Separate tutorials are provided for each Multi-Component Analysis sub-module, where the examples use experimental SANS CV data directly or incorporate parameters derived from these data. The tutorials contain descriptions of the input variables and the output results, including links to download the input and output files, where applicable. Screenshots at every step help the user complete the tutorial and references are provided for further reading. As the documentation expands, these tutorials will be updated with new features, refined instructions and additional real-data examples, ensuring they remain a practical and up-to-date guide for the *SASSIE-web* community.

5. Conclusions

The new *SASSIE-web* SANS CV modules, integrated with existing *SASSIE-web* modules, allow users to plan a SANS CV experiment and analyze SANS CV data from two-component biological complexes. Using this two-pronged approach, the new modules aim to reduce the time and resources required for complex CV experiments while increasing the chances of successful outcomes. For both experienced researchers and those navigating the steep learning curve of SANS CV, this integrated platform represents a significant step forward in accessibility to this specialized type of small-angle scattering experiment.

Acknowledgements

The authors thank Susana Teixeira and Nick Clark for testing the software and providing valuable feedback. Certain commercial equipment, instruments, materials, suppliers or software are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Funding information

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award No. R01GM120600 (to EB). This work used Jetstream2 at Indiana University through

allocation MCB170057 (to EB) from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by US National Science Foundation (NSF) grants No. 2138259, No. 2138286, No. 2138307, No. 2137603 and No. 2138296. This work benefited from the Collaborative Computational Project for Small Angle Scattering (*CCP-SAS*) software developed through a joint EPSRC (EP/K039121/1) and NSF (CHE-1265821) grant.

References

- Brookes, E. H., Anjum, N., Curtis, J. E., Marru, S., Singh, R. & Pierce, M. (2015). *Concurrency Comput. Pract. Exp.* **27**, 4292–4303.
- Curtis, J. E., Raghunandan, S., Nanda, H. & Krueger, S. (2012). *Comput. Phys. Commun.* **183**, 382–389.
- Engelman, D. & Moore, P. B. (1975). *Q. Rev. Biophys.* **4**, 219–241.
- Glatter, O. (1977). *J. Appl. Cryst.* **10**, 415–421.
- Guinier, A. & Fournet, G. (1955). *Small-angle scattering of X-rays*. New York: John Wiley & Sons.
- Ibel, K. & Stuhmann, H. B. (1975). *J. Mol. Biol.* **93**, 255–265.
- Krueger, S. (2017). *Biological small angle scattering: techniques, strategies and tips*, Advances in experimental medicine and biology, Vol. 1009, edited by B. Chaudhuri, I. G. Muñoz, S. Qian & V. S. Urban, pp. 65–85. Springer Singapore.
- Krueger, S. (2022a). *Methods Enzymol.* **677**, 127–155.
- Krueger, S. (2022b). *Curr. Opin. Struct. Biol.* **74**, 102375.
- Mahieu, E. & Gabel, F. (2018). *Acta Cryst. D* **74**, 715–726.
- Moore, P. B. (1982). *Methods of experimental physics*, Vol. 20, *Biophysics*, edited by G. Ehrenstein & H. Lecar, pp. 337–390. New York: Academic Press.
- Olah, G. A., Rokop, S. E., Wang, C. A., Blechner, S. L. & Trehwella, J. (1994). *Biochemistry* **33**, 8233–8239.
- Perkins, S. J., Wright, D. W., Zhang, H., Brookes, E. H., Chen, J., Irving, T. C., Krueger, S., Barlow, D. J., Edler, K. J., Scott, D. J., Terrill, N. J., King, S. M., Butler, P. D. & Curtis, J. E. (2016). *J. Appl. Cryst.* **49**, 1861–1875.
- Plotly Technologies (2015). *Plotly*. Plotly Technologies Inc., Montréal, Quebec, Canada.
- Sarachan, K. L., Curtis, J. E. & Krueger, S. (2013). *J. Appl. Cryst.* **46**, 1889–1893.
- Semenyuk, A. V. & Svergun, D. I. (1991). *J. Appl. Cryst.* **24**, 537–540.
- Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.
- Trehwella, J. (2022). *Structure* **30**, 15–23.
- Trehwella, J. (2023). *Methods Enzymol.* **678**, 1–22.
- Trehwella, J., Duff, A. P., Durand, D., Gabel, F., Guss, J. M., Hendrickson, W. A., Hura, G. L., Jacques, D. A., Kirby, N. M., Kwan, A. H., Pérez, J., Pollack, L., Ryan, T. M., Sali, A., Schneidman-Duhovny, D., Schwede, T., Svergun, D. I., Sugiyama, M., Tainer, J. A., Vachette, P., Westbrook, J. & Whitten, A. E. (2017). *Acta Cryst. D* **73**, 710–728.
- Whitten, A. E., Cai, S. & Trehwella, J. (2008). *J. Appl. Cryst.* **41**, 222–226.
- Whitten, A. E. & Jeffries, C. M. (2023). *Methods Enzymol.* **678**, 55–96.
- Zaccai, N. R., Sandlin, C. W., Hoopes, J. T., Curtis, J. E., Fleming, P. J., Fleming, K. G. & Krueger, S. (2016). *Methods Enzymol.* **566**, 159–210.