Journal of
Applied Crystallography

ISSN 0021-8898

Received 10 August 2013
Accepted 17 September 2013

Small-angle scattering contrast calculator for protein and nucleic acid complexes in solution

Kathryn L. Sarachan, Joseph E. Curtis* and Susan Krueger*

NIST Center for Neutron Research, NIST, 100 Bureau Drive, Stop 6102, Gaithersburg, MD 20899-6102, USA. Correspondence e-mail: joseph.curtis@nist.gov, susan.krueger@nist.gov

Small-angle neutron scattering (SANS) with contrast variation can provide useful information about the structure and disposition of two or more chemically distinct components within a complex. The *SASSIE Contrast Calculator* (*SCC*) is a new software tool designed to assist in planning SANS experiments with contrast variation on protein and nucleic acid complexes. On the basis of the primary sequence and deuteration level of each protein or nucleic acid component, the *SCC* calculates and plots $I(0)$, contrast and scattering length densities; since SANS experiments often complement small-angle X-ray scattering studies, the program provides both neutron and X-ray parameters. The *SCC* is run as an integrated component of *SASSIE* [Curtis, Raghunandan, Nanda & Krueger (2012). *Comput. Phys. Commun.* **183**, 382–389], a software suite for atomistic modeling of ensembles of structures consistent with scattering data.

1. Introduction

Small-angle scattering (SAS) techniques based on the diffraction of X-rays or neutrons (SAXS/SANS) from homogeneous biological samples in solution have become established methods employed by structural biologists (Petoukhov & Svergun, 2007; Whitten & Trewhella, 2009) to provide low-resolution information about the structures of biomacromolecules and their assemblies. SAS fills a role in structural biology complementary to that of high-resolution techniques such as X-ray crystallography and NMR, especially for arbitrarily large or dynamic systems refractory to more traditional approaches. In recent years, as tandem advances in data collection, instrumentation and computing power have enabled the routine reconstruction of three-dimensional molecular envelopes of scattering particles, the popularity of the technique has soared (Chacón *et al.*, 1998; Hura *et al.*, 2009).

SANS offers a particular advantage over SAXS in the study of composite systems containing components with different average scattering length densities, such as protein–nucleic acid complexes or protein–protein complexes in which one component has been labeled with deuterium. Capitalizing on the large difference in neutron scattering length between hydrogen and deuterium, a SANS contrast variation experiment can determine the specific scattering contributions of the constituents of a complex by systematically varying the solvent H_2O:D_2O ratio to 'match' the scattering length density of each component. This contrast matching provides a valuable tool for examining the structure and disposition of each component in the context of the complex as a whole (Ankner *et al.*, 2013; Heller, 2010).

For many years, the experimental parameters for a given SANS contrast variation experiment were intuited on the basis of previous experience or, at best, calculated using homemade computer programs not widely accessible. Recently, a web-based program, *MULCh*, became available that calculates SAXS and SANS scattering length densities and analyzes SANS contrast variation data for two-component systems that may include deuterated proteins, DNA or RNA; it also considers non-water components in the solvent (Whitten *et al.*, 2008). It is a very useful tool for some aspects of experiment planning and data analysis. However, since *MULCh* can only be used for two-component systems, it cannot be used to plan SANS contrast variation experiments on more complicated systems. Also, since forward scattering $I(0)$ values are not provided, it is difficult to predict the signal-to-noise ratio of the SANS data for complexes in various H_2O:D_2O solvents.

Here we present the *SASSIE Contrast Calculator* (*SCC*), a software tool designed to assist in experiment planning for small-angle neutron scattering experiments with contrast variation for an arbitrary number of protein and nucleic acid (DNA and RNA) scattering species. The *SCC* requires as input only primary sequence data, accepted as either PDB or FASTA files, and deuteration levels in order to calculate and plot $I(0)$, scattering length densities and contrast. The *SCC* is an integrated component of the *SASSIE* software suite (Curtis *et al.*, 2012).

2. Theory

The X-ray and neutron scattering length densities (SLDs) for proteins and nucleic acids are calculated directly from the atomic composition of each amino acid residue and/or nucleotide in the molecule. For X-rays, the scattering length of a molecule can be written as

$$b_{\text{X-ray}} = \sum_i r_e Z_i, \tag{1}$$

where r_e is the classical (Thompson) electron radius and Z_i is the atomic number of the ith atom in the molecule. Under this formalism, for example, the X-ray scattering length of water is $r_e(2Z_H + Z_O)$. Similarly, the X-ray scattering length of a protein or nucleic acid molecule can be calculated simply by adding the atomic numbers of all of the atoms in the molecule and multiplying the result by r_e.

Since neutrons interact with the nucleus of an atom, neutron scattering lengths do not scale with increasing atomic number.

Furthermore, the different neutron scattering lengths of hydrogen and deuterium must be taken into account when calculating neutron scattering lengths of proteins or nucleic acids when they are partially deuterated or dissolved in a solution containing some fraction of D_2O. Thus, the neutron scattering length of the solution is

$$b_{\rm soln} = b_{\rm H2O} + f_{\rm D2O}\left(b_{\rm D2O} - b_{\rm H2O}\right), \quad (2)$$

where $b_{\rm H2O}$ and $b_{\rm D2O}$ are the neutron scattering lengths of H_2O and D_2O, respectively, and $f_{\rm D2O}$ is the fraction of D_2O in the solution. The neutron scattering length of a protein or nucleic acid molecule can be calculated from

$$b_{\rm neutron} = \sum_i b_i + (n_{\rm H} - n_{\rm Hex}) f_{\rm deut}(b_{\rm D} - b_{\rm H}) + n_{\rm Hex} f_{\rm exch} f_{\rm D2O}(b_{\rm D} - b_{\rm H}), \quad (3)$$

where b_i is the neutron scattering length of the ith atom in the molecule, $n_{\rm H}$ is the total number of H atoms in the molecule, $n_{\rm Hex}$ is the total number of exchangeable H atoms in the molecule, $f_{\rm deut}$ is the fraction of deuteration in the molecule, $b_{\rm D}$ and $b_{\rm H}$ are the neutron scattering lengths of deuterium and hydrogen, respectively, and $f_{\rm exch}$ is the fraction of exchangeable H atoms that exchange with D in the solution.

Both X-ray and neutron scattering length densities (ρ) are found by dividing the scattering lengths by the volume, such that

$$\rho_{\rm protein} = \frac{b_{\rm protein}}{V_{\rm protein}},\ \rho_{\rm nucleic\ acid} = \frac{b_{\rm nucleic\ acid}}{V_{\rm nucleic\ acid}} \text{ and } \rho_{\rm soln} = \frac{b_{\rm soln}}{V_{\rm H2O}}. \quad (4)$$

For proteins and nucleic acids, the molecular volume is estimated from the corresponding partial specific volumes using the relation

$$V = \overline{v} M_{\rm w}/N_{\rm A}, \quad (5)$$

where $\overline{v}$ is the partial specific volume of the protein or nucleic acid, $M_{\rm w}$ is its molecular weight and $N_{\rm A}$ is Avogadro's number. When deuterium is present in the molecule or solution, the calculation of $M_{\rm w}$ is adjusted such that

$$M_{\rm w} = \sum_j M_{{\rm w}j} + f_{\rm deut}(n_{\rm H} - n_{\rm Hex}) + n_{\rm Hex} f_{\rm exch} f_{\rm D2O}, \quad (6)$$

where j refers to each individual component, *i.e.* protein, DNA or RNA.

The difference in scattering length densities between the protein or nucleic acid and the solution is the contrast, $\Delta\rho$,

$$\Delta\rho = \rho - \rho_{\rm soln} = \frac{bN_{\rm A}}{\overline{v}M_{\rm w}} - \frac{b_{\rm soln}}{V_{\rm H2O}}. \quad (7)$$

For a complex consisting of separate components with different scattering length densities relative to the solvent, such as proteins and nucleic acids, the contrast is weighted by the mass fraction of each component such that

$$\Delta\rho = \sum_j f_j \Delta\rho_j, \quad (8)$$

where j refers to the individual components and f_j is given by $M_{{\rm w}j}/\sum_j M_{{\rm w}j}$.

The forward scattering $I(0)$ is calculated on an absolute scale, in units of $\rm cm^{-1}$, from the equation

$$I(0) = n(\Delta\rho)^2 V^2, \quad (9)$$

where n is the number density, in units of $\rm cm^{-3}$. n is related to the concentration, c, by the relation

$$n = cN_{\rm A}/M_{\rm w}. \quad (10)$$

Combining equation (9) with equations (5), (8) and (10) results in

$$I(0) = (cM_{\rm w}/N_{\rm A})\left(\sum_j f_j \Delta\rho_j \overline{v}_j\right)^2, \quad (11)$$

where $M_{\rm w}$ now refers to the total molecular weight of the entire complex. It is important to note that, in the case of neutrons, this $I(0)$ value is for the coherent scattering component only. It does not include the contribution from incoherent scattering from H or D atoms in the complex or the solvent. For example, the neutron incoherent scattering contribution to $I(0)$ is $\sim$0.07 $\rm cm^{-1}$ for D_2O and 1.0 $\rm cm^{-1}$ for H_2O.

3. Description of program

The *SASSIE Contrast Calculator* evaluates the neutron and X-ray SLDs, contrast, and $I(0)$ values for a single molecule or complex based only on the primary sequences of its protein and/or nucleic acid constituents. The input files can be in either FASTA or PDB format, and there is no limit to the number of protein and/or nucleic acids that may make up the complex. The neutron values are calculated as a function of the D_2O fraction [$f_{\rm D2O}$ from equation (2)] in the solvent; the fraction of deuteration for the proteins and/or nucleic acids is also considered in the calculation, as are, optionally, non-water solvent components. The SLD, contrast and $I(0)$ information are output in separate files containing the input information as well as the output parameters for both X-rays and neutrons. The *SCC* is written in Python and is fully integrated with the *SASSIE* software package (Curtis *et al.*, 2012), where it can be found as a module under the Tools menu in the *SASSIE* graphical user interface (GUI). A tutorial with step-by-step examples and screenshots is available with the software (see §5).

3.1. Program GUI and inputs

The *SCC* GUI window (Figs. S1 and S6 in the supplemental material[1]) is divided into three parts. The top of the window shows the User Input Section, which contains input fields for the project directory and output file names as well as the number of input files (FASTA sequences or PDB files). An Input File Information dialog box is used to record the input file information such as file name, number of units and fraction of deuteration. The parameter indicating the number of units can be greater than one to account for multiple occurrences of identical chains or subunits, and it must be entered correctly in order that the $I(0)$ value be calculated with the appropriate $M_{\rm w}$ for the complex. A checkbox is provided to indicate if the input file is in FASTA format; otherwise, it is assumed to be in PDB format. Radio buttons indicate whether the file represents a protein, DNA or RNA component.

The middle of the window contains Optional Input, which includes input fields for the total solute concentration (default = 1 mg $\rm ml^{-1}$), the step in solvent percent D_2O (default = 5% increments in percent D_2O from 0 to 100%), the fraction of exchangeable H atoms in the protein and nucleic acid components (default = 0.95 and 1.0, respectively) that actually exchange, and the number of non-water solvent components (optional; default = 0). Information regarding non-water solvent components is entered using a separate Solvent Component Information dialog box that accepts the atomic formula and concentration in mol $\rm l^{-1}$ (M) of each non-water component.

[1] Supplementary material, featuring two examples illustrating the use of the *SASSIE Contrast Calculator*, with accompanying figures, is available from the IUCr electronic archives (Reference: FS5057). Services for accessing this material are described at the back of the journal.

The bottom of the window contains a message box where the status of the calculations and other information such as component match points are provided. The 'Run Contrast Calculator Program' button below the status box is used to start the calculations once all necessary parameters have been entered. Finally, a progress bar at the bottom of the window tracks the progression of the contrast calculations.

3.2. Program flow

Each scattering component of the complex under consideration is defined by both molecule type and fraction of deuteration. Within each given molecule type (protein, DNA or RNA), the input molecules are binned according to their respective fractions of deuteration [$f_{\rm deut}$ in equation (3)]. Scattering lengths are combined for all files of a given molecule type that have the same fraction of deuteration. SLDs and contrasts are then calculated for each component. This step is accomplished differently for X-ray and neutron SLD and contrast calculations; in the former case, the X-ray scattering lengths for hydrogen and deuterium are the same, meaning that only the molecule type is important for the X-ray calculations. Finally, the contrast and $I(0)$ values are calculated for the entire complex, weighted by the mass fraction and, in the case of $I(0)$, the partial specific volume of each component.

3.2.1. Solvent parameters. Non-water solvent components such as salts, sugars, cryoprotectants such as polyethylene glycols, and other buffer components are treated in the same manner as described by Whitten *et al.* (2008); their influence on the scattering length density of the solvent is presumed to be a result of excluded volume effects. While neutron solvent scattering length densities are usually not significantly affected by non-water solvent components at mM concentration levels, the effect on X-ray scattering length densities may be higher depending on the total Z value of the components, and non-water solvent components at molar concentration levels should always be included in the calculation of solution parameters.

3.2.2. Protein and nucleic acid parameters. Tables of amino acid residues and nucleotides contain information on their $M_{\rm w}$, volume, X-ray and neutron scattering lengths, total number of H atoms, and number of exchangeable H atoms. The atomic formulae used to calculate these parameters assume that the amino acid residues and nucleotides are embedded in a chain in the manner of Jacrot (1976). The partial specific volume for protein components is taken to be 0.73 cm^3 (g m)$^{-1}$ (Voss & Gerstein, 2005). Partial specific volumes for DNA are dependent on pH values and salt and occupy a range between 0.55 and 0.59 cm^3 (g m)$^{-1}$ (Hearst, 1962). For RNA, the partial specific volume is larger for double-stranded RNA than for single-stranded RNA, and values range from 0.47 to 0.55 cm^3 (g m)$^{-1}$ (Chien *et al.*, 2004). Here, the value used for both DNA and RNA in the *SCC* is taken to be 0.56 cm^3 (g m)$^{-1}$. The fractions of exchangeable protein and nucleic acid hydrogen atoms that exchange with deuterium in the solvent [$f_{\rm exch}$ in equation (3)] are adjustable parameters with default values of, respectively, 0.95 for proteins and 1.0 for nucleic acids. The protein default value is based on the empirical observation that values between 0.85 and 0.95 do not appreciably change calculated scattering lengths, as would be expected given that $f_{\rm exch}$ is only present in the third term in equation (3).

For the purpose of calculating X-ray SLDs, the scattering lengths and $M_{\rm w}$ values for all protein and nucleic acid (RNA and DNA) components are combined since the X-ray scattering lengths for hydrogen and deuterium are the same and the $M_{\rm w}$ values have already been adjusted for any deuterium in the molecule. The X-ray SLDs are calculated using equations (1), (4), (5) and (6), and the X-ray contrast is then calculated for each of the two components using equation (7). Finally, the combined X-ray contrast for the entire complex is calculated using equation (8), noting the difference in partial specific volumes for proteins and nucleic acids. $I(0)$ for the complex is calculated using equation (11).

Neutron scattering lengths and $M_{\rm w}$ values are calculated separately for each component within the three molecule types using equation (3), and corresponding neutron SLDs are calculated using equations (4), (5) and (6). The neutron contrasts are then calculated for each component as a function of D_2O fraction in the solvent using equations (2) and (7), and the combined neutron contrast for the entire complex is weighted by mass fraction as in equation (8). $I(0)$ is calculated using equation (11). Neutron contrast match points are calculated for each component and for the entire complex by performing a linear fit to the contrast *versus* solvent $f_{\rm D2O}$ in each case. The match point for the entire complex can also be determined from a linear fit to $I(0)^{1/2}$ *versus* $f_{\rm D2O}$.

3.3. Program outputs

Once the calculations are finished, the match points of each of the components, as well as that of the total complex, are written to the message box, together with date and time information and the names of the output files along with their location. If the complex contains fewer than three components, then plots of the neutron SLD *versus* $f_{\rm D2O}$, the contrast *versus* $f_{\rm D2O}$, $I(0)$ *versus* $f_{\rm D2O}$ and $I(0)^{1/2}$ *versus* $f_{\rm D2O}$ are also output. In addition, three output files are generated with file names that correspond to the user-supplied output file name with added extensions: `file_sld.txt` contains the X-ray SLD, the neutron match points and the neutron SLDs as a function of the D_2O fraction in the solvent; `file_contrast.txt` contains the X-ray contrast, the neutron match points and the neutron contrast as a function of the fraction of solvent D_2O; and `file_izero.txt` contains the total concentration of the complex, the X-ray $I(0)$ in cm^{-1}, the neutron match points, and the $M_{\rm w}$ and neutron $I(0)$ values in cm^{-1} as a function of the fraction of solvent D_2O. All three files specify the time and date, the input filenames with the number of units in parentheses, the solvent information, and the fraction of exchangeable H atoms for the protein and nucleic acid components.

3.4. Examples

Examples of the application of the *SCC* to a protein–protein system (a complex comprising a skp trimer bound to OmpA in which the latter protein is partially deuterated) and to a protein–DNA system [two synapsis-deficient Cre(A36V) recombinase molecules bound to a single *loxP* site] can be found in the supplemental material. The skp–OmpA and Cre(A36V) primary sequences are described elsewhere (Walton & Sousa, 2004; Ghosh *et al.*, 2007).

4. Experiment planning

The *SCC* is designed to calculate parameters useful in planning SANS contrast variation experiments. For example, one might consider a two-component complex consisting of protein and deuterated protein or protein and DNA. The neutron scattering at a given contrast can be written as

$$I(q) = \Delta\rho_1^2 I_{11}(q) + \Delta\rho_1\Delta\rho_2 I_{12}(q) + \Delta\rho_2^2 I_{22}(q), \qquad (12)$$

where $I_{11}(q)$ and $I_{22}(q)$ are the scattering curves from each component and $I_{12}(q)$ is the interference scattering curve between the two components. These scattering curves are the three unknowns in the equation. Thus, $I(q)$ must be measured at a minimum of three

contrasts in order to solve a set of simultaneous equations that will provide these individual scattering curves. Typically, four to five contrast points are measured for redundancy. As this requires quite a bit of material and preparation, judicious planning prior to the contrast variation experiment is prudent.

The *SCC* provides SLD and contrast information for each component that can be used, together with the calculated contrast match points, to determine optimum contrasts for the experiment. In addition, the $I(0)$ *versus* f_{D2O} values are extremely useful for determining whether a reasonable signal will be obtained at those contrasts for a given concentration. The experimental neutron $I(0)$ value comprises both a coherent and an incoherent contribution, as

$$I(0)_{exp} = I(0)_{coh} + I_{incoh}, \quad (13)$$

where $I(0)_{exp}$ is the experimental $I(0)$, $I(0)_{coh}$ is the coherent contribution to that value and I_{incoh} the incoherent scattering. The $I(0)$ value calculated by the *SCC* corresponds only to $I(0)_{coh}$, and experimental planning must take into account the likely effect of incoherent scattering on the measurable signal. While the incoherent contribution from the complex itself can usually be neglected in the case of dilute solutions, the solvent contribution can be substantial. For example, since the incoherent component for H_2O is more than an order of magnitude higher than that for D_2O, the signal-to-noise ratio in the measured data will be lower for a complex in H_2O solvent, even if the $I(0)$ value calculated by the *SCC* is the same as that in D_2O solvent.

Fig. 1 shows the neutron scattering length density *versus* f_{D2O} plots for the skp–OmpA protein–deuterated protein and the Cre–*loxP* protein–DNA complexes that are described in detail in the supplemental material. The solvent curve intersects those of the components and the complex at the corresponding match points. The curves for the skp–OmpA complex (Fig. 1*a*) show that both the skp and the 50% deuterated OmpA match points are well separated from that of the complex as a whole. It should, therefore, be possible to make measurements at contrast points close to those of the two components, *i.e.* 44% D_2O and 80% D_2O, and a contrast variation experiment might consist of measurements at 0% D_2O, 44% D_2O, 80% D_2O and 100% D_2O to solve for the scattering curves in equation (12). For redundancy, a contrast point can be added near 20% D_2O as well. Table 1 shows the calculated M_w and neutron $I(0)$ values *versus* D_2O fraction for the skp–OmpA complex. The table reveals that the $I(0)$ value for the complex at 1 mg ml^{-1} in 44% D_2O is very low and would be difficult to measure, especially given that the incoherent scattering contribution would still be significant under these conditions. The $I(0)$ value at 80% D_2O is higher and the contribution from incoherent scattering smaller, but it still represents a difficult measurement. Either the concentration of the complex must be increased or alternative contrast points further away from these match points must be chosen. Since it is desirable to measure as close to the individual component match points as possible in order to obtain the most optimal solutions to equation (12), it would be advisable in this case to measure the complex at a higher concentration if at all possible.

Table 1
Calculated M_w and neutron $I(0)$ *versus* solvent D_2O fraction for a 1 mg ml^{-1} protein–protein complex comprising a skp trimer bound to 50% deuterated OmpA.

D_2O fraction	skp M_w (kDa)	OmpA (50%D) M_w (kDa)	Complex M_w (kDa)	$I(0)$ (cm^{-1})
0.0	49.91	36.22	86.13	0.086
0.1	50.00	36.27	86.27	0.059
0.2	50.08	36.33	86.41	0.038
0.3	50.17	36.38	86.55	0.021
0.4	50.26	36.43	86.69	0.009
0.5	50.35	36.49	86.74	0.002
0.6	50.44	36.54	86.98	0.000
0.7	50.53	36.59	87.12	0.003
0.8	50.62	36.65	87.27	0.010
0.9	50.70	36.70	87.40	0.022
1.0	50.79	36.75	87.54	0.040

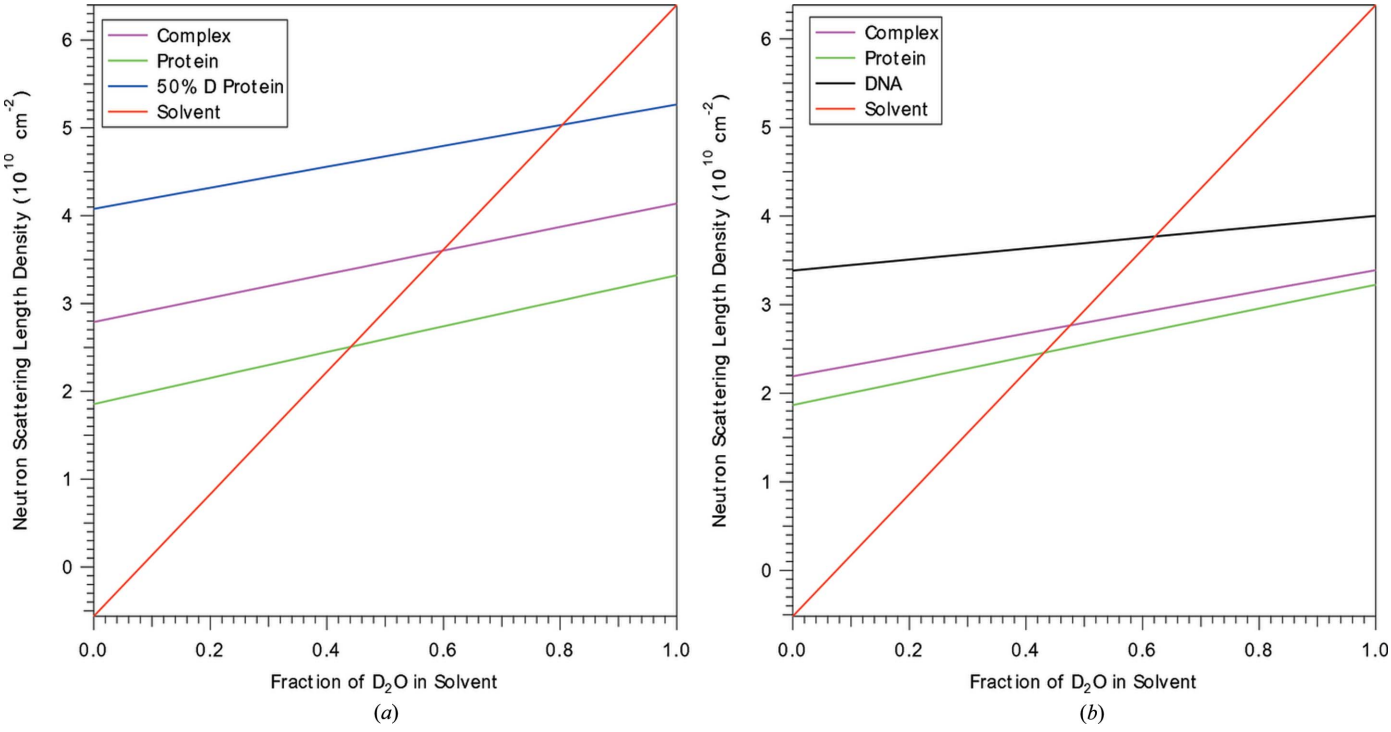

Figure 1
Neutron scattering length density *versus* D_2O fraction for (*a*) a skp–OmpA protein–deuterated protein complex and (*b*) a Cre–*loxP* protein–DNA complex.

Table 2
Calculated M_w and neutron $I(0)$ *versus* solvent D_2O fraction for a 1 mg ml^{-1} protein–DNA Cre(A36V)–*loxP* complex.

D_2O fraction	Cre(A36V) M_w (kDa)	*loxP* M_w (kDa)	Complex M_w (kDa)	$I(0)$ (cm^{-1})
0.0	77.00	20.99	97.99	0.055
0.1	77.12	21.00	98.12	0.034
0.2	77.25	21.01	98.26	0.018
0.3	77.37	21.02	98.39	0.007
0.4	77.50	21.04	98.54	0.001
0.5	77.63	21.05	98.68	0.000
0.6	77.75	21.06	98.81	0.004
0.7	77.88	21.07	98.95	0.014
0.8	78.01	21.08	99.09	0.028
0.9	78.13	21.10	99.23	0.048
1.0	78.26	21.11	99.37	0.073

Realistically, a total concentration of at least 5 mg ml^{-1} for the complex would produce much better results.

On the other hand, the curves for the Cre–*loxP* complex (Fig. 1*b*) show that only the match point for the DNA component is well separated from that of the complex. Thus, it will not be possible to make measurements at the Cre match point (*i.e.* 43% D_2O) and it will be more difficult to solve equation (12) for the DNA component. Table 2 shows the calculated M_w and neutron $I(0)$ values *versus* D_2O fraction for the Cre–*loxP* complex and indicates that the $I(0)$ values at 20% D_2O and 70% D_2O (*i.e.* the DNA match point) would possibly be measurable. The incoherent scattering contribution would be more problematic at 20% D_2O than at 70% D_2O. A contrast variation experiment might consist of measurements at 0% D_2O, 20% D_2O, 70% D_2O and 100% D_2O; again, a concentration for the complex higher than 1 mg ml^{-1} would be desirable. If possible, it would also be desirable to add a fifth contrast point near 30% D_2O in order to try to optimize the solutions to equation (12). Even at a total concentration of 5 mg ml^{-1} or higher, the solution for the *loxP* DNA component may lack sufficient signal to be usable because the Cre match point and the complex match point are too close to each other. However, even solving for just the protein component in the complex can be useful, especially if the result is to be compared with that from the protein alone in solution to determine if the protein changes conformation upon DNA binding.

Table 3 compares the neutron SLD values of the solvent when it contains 0.5 *M* NaCl with those where the solvent contains no salt. The presence of salt increases the solvent neutron SLD value by about 10% in 0% D_2O and decreases it by about 1% in 100% D_2O. Higher salt concentrations would contribute to a larger degree; smaller amounts of different non-water solvent components might also have a large effect. It is therefore useful and convenient to include non-water solvent components in the calculation of the contrast parameters.

Table 3
Neutron scattering length density (SLD) values for solvent without and with 0.5 *M* NaCl.

D_2O fraction	Solvent SLD (10^{10} cm^{-2})	Solvent SLD (0.5 *M* NaCl) (10^{10} cm^{-2})
0.0	−0.561	−0.517
0.1	0.135	0.173
0.2	0.831	0.863
0.3	1.527	1.554
0.4	2.223	2.244
0.5	2.920	2.934
0.6	3.616	3.624
0.7	4.312	4.314
0.8	5.008	5.004
0.9	5.704	5.694
1.0	6.400	6.384

5. Distribution and tutorial

The *SCC* is distributed within the *SASSIE* software suite and currently supports Linux and MacOS X operating systems (Curtis *et al.*, 2012). The software and tutorials can be obtained from http://www.smallangles.net/sassie.

6. Conclusions and future directions

The *SASSIE Contrast Calculator* (*SCC*) is a useful tool to aid in the planning of SAS experiments on biological complexes consisting of proteins and/or nucleic acids, with a particular application toward determining optimal parameters for neutron contrast variation measurements. Currently, plans exist to expand the capabilities of the software beyond those presented here. One important anticipated feature will be the prediction, when PDB-formatted input files are available, of complete SAXS and SANS $I(q)$ *versus* q scattering curves on an absolute scale for a given concentration of the biological complex in solution. Additionally, for neutrons, it will be possible to calculate the scattering curves at a chosen solvent D_2O fraction, perhaps on the basis of an initial run of the *SCC* over a wider range of values between 0 and 100%. An estimate of the incoherent scattering level for the chosen contrast points will be provided as another planning aid. Finally, a Stuhrmann analysis (Ibel, 1975) will be performed on the calculated scattering curves to provide a prediction of how the radius of gyration might change with contrast for a given biological complex. These additional features will enhance the ability of the *SASSIE Contrast Calculator* to provide a robust platform for use in the design of small-angle neutron and X-ray scattering experiments.

This material is based upon work supported by the National Science Foundation under CHE-1265817. KLS gratefully acknowledges support from a NIST National Research Council (NRC) postdoctoral research associateship.

References

Ankner, J. F., Heller, W. T., Herwig, K. W., Meilleur, F. & Myles, D. A. A. (2013). *Current Protocols in Protein Science*, edited by J. E. Coligan, B. M. Dunn, D. W. Speicher & P. T. Wingfield. Hoboken: John Wiley and Sons.
Chacón, P., Morán, F., Díaz, J. F., Pantos, E. & Andreu, J. M. (1998). *Biophys. J.* **74**, 2760–2775.
Chien, C. Y., Xu, Y., Xiao, R., Aramini, J. M., Sahasrabudhe, P. V., Krug, R. M. & Montelione, G. T. (2004). *Biochemistry*, **43**, 1950–1962.
Curtis, J. E., Raghunandan, S., Nanda, H. & Krueger, S. (2012). *Comput. Phys. Commun.* **183**, 382–389.
Ghosh, K., Guo, F. & Van Duyne, G. D. (2007). *J. Biol. Chem.* **282**, 24004–24016.
Hearst, J. E. (1962). *J. Mol. Biol.* **4**, 415–417.
Heller, W. T. (2010). *Acta Cryst.* D**66**, 1213–1217.
Hura, G. L., Menon, A. L., Hammel, M., Rambo, R. P., Poole, F. L., Tsutakawa, S. E., Jenney, F. E., Classen, S., Frankel, K. A., Hopkins, R. C., Yang, S. J., Scott, J. W., Dillard, B. D., Adams, M. W. & Tainer, J. A. (2009). *Nat. Methods*, **6**, 606–612.
Ibel, K. (1975). *J. Mol. Biol.* **93**, 255–265.
Jacrot, B. (1976). *Rep. Prog. Phys.* **39**, 911–953.
Petoukhov, M. V. & Svergun, D. I. (2007). *Curr. Opin. Struct. Biol.* **17**, 562–571.
Voss, N. R. & Gerstein, M. (2005). *J. Mol. Biol.* **346**, 477–492.
Walton, T. A. & Sousa, M. C. (2004). *Mol. Cell*, **15**, 367–374.
Whitten, A. E., Cai, S. & Trewhella, J. (2008). *J. Appl. Cryst.* **41**, 222–226.
Whitten, A. E. & Trewhella, J. (2009). *Methods Mol. Biol.* **544**, 307–323.