Preparing monodisperse macromolecular samples for successful biological small-angle X-ray and neutron-scattering experiments

Cy M Jeffries[1], Melissa A Graewert[1], Clément E Blanchet[1], David B Langley[2,3], Andrew E Whitten[4] & Dmitri I Svergun[1]

[1]European Molecular Biology Laboratory (EMBL) Hamburg Outstation, DESY, Hamburg, Germany. [2]Victor Chang Cardiac Research Institute, Darlinghurst, New South Wales, Australia. [3]Garvan Institute, Darlinghurst, New South Wales, Australia. [4]Australian Centre for Neutron Science, ANSTO, Lucas Heights, New South Wales, Australia. Correspondence should be addressed to C.M.J. (cy.jeffries@embl-hamburg.de).

Published online 6 October 2016; doi:10.1038/nprot.2016.113

Small-angle X-ray scattering (SAXS) and small-angle neutron scattering (SANS) are techniques used to extract structural parameters and determine the overall structures and shapes of biological macromolecules, complexes and assemblies in solution. The scattering intensities measured from a sample contain contributions from all atoms within the illuminated sample volume, including the solvent and buffer components, as well as the macromolecules of interest. To obtain structural information, it is essential to prepare an exactly matched solvent blank so that background scattering contributions can be accurately subtracted from the sample scattering to obtain the net scattering from the macromolecules in the sample. In addition, sample heterogeneity caused by contaminants, aggregates, mismatched solvents, radiation damage or other factors can severely influence and complicate data analysis, so it is essential that the samples be pure and monodisperse for the duration of the experiment. This protocol outlines the basic physics of SAXS and SANS, and it reveals how the underlying conceptual principles of the techniques ultimately 'translate' into practical laboratory guidance for the production of samples of sufficiently high quality for scattering experiments. The procedure describes how to prepare and characterize protein and nucleic acid samples for both SAXS and SANS using gel electrophoresis, size-exclusion chromatography (SEC) and light scattering. Also included are procedures that are specific to X-rays (in-line SEC–SAXS) and neutrons, specifically preparing samples for contrast matching or variation experiments and deuterium labeling of proteins.

INTRODUCTION

Modern systems structural biology is faced with enormous challenges in deciphering the complexity of interconnected macromolecular networks and how these networks mediate molecular-level communication to affect cellular responses. High-resolution structure determination methods, such as X-ray crystallography, NMR spectroscopy and, more recently, high-resolution electron microscopy (EM) are exceptional for uncovering the atomic details of proteins and other macromolecules. However, it becomes increasingly difficult using high-resolution techniques to assess the conformational responses of macromolecules, complexes and assemblies in different sample environments. Continued advances in instrumentation and software, and the development of automated methods for data collection, analysis and modeling have launched small-angle scattering (SAS) using X-rays (SAXS) or neutrons (SANS) into the structural biological mainstream[1–8]. The appeal of SAS for structural biologists is that it can be applied to the analysis of diverse macromolecular systems—directly in solution—that span a broad molecular weight (MW) range, from a few kilodaltons to megadaltons, across a seemingly endless array of sample environments[9–12]. Global structural parameters—for example, the radius of gyration, R_g; maximum particle dimension, D_{max}; and the distribution of distances within a particle (relating to the volume and structure)—can be quickly extracted from the data[2,6]. In addition, it is now routine to obtain low-resolution 3D-spatial representations of macromolecules using SAS[4,13–17]. Modeling these structures can be achieved using combination(s) of *ab initio* methods[18–20], where no prior assumptions are made, or using atomistic or rigid-body models[4,21] derived from X-ray crystallography, NMR, EM and homology modeling (i.e., hybrid methods)[22]. Importantly, as solution environments can be controlled, SAS is extremely useful for probing the structural responses of macromolecules to changing sample conditions[23]. Ensemble states[24–28]—as examined, for example, in the study of intrinsically disordered proteins[29,30] and the formation of complexes[31] or assemblies in real time[32–34]—can be evaluated using SAS; such evaluations are otherwise difficult to achieve using high-resolution methods.

One drawback of SAS is that it is difficult to prove with certainty that a measured scattering profile is, in fact, derived from a target of interest. All matter has the potential to scatter radiation (X-rays and neutrons), and thus all atoms comprising a sample—macromolecules, water, buffer components, macromolecules, the sample container and so on—will each contribute to the measured scattering intensities. Fundamentally, the success of any experiment will rely on the production of well-characterized, high-quality samples[35,36], combined with an accurate understanding of, and correction for, any background scattering contributions. Consequently, maintaining sample quality for SAXS and SANS is challenging partly because of how the physics of the two techniques relates to the properties of a sample.

For the structural biologist at the laboratory bench who is interested in applying SAXS or SANS to interrogate the structures of macromolecules in solution, a great deal of the physics describing the two techniques can be difficult to translate into a procedure for sample preparation. In practice, all that is required

is an understanding of a handful of concepts that help define what practical steps are necessary to produce quality samples. Advanced and detailed explanations of the physics and mathematics of SAS—that become increasingly relevant when analysing datasets or for the design of experiments—can be found in the texts by Glatter and Kratky[37], Feigin and Svergun[38] and the more recent Svergun *et al.*[39]. Additional protocols for SAS data acquisition, basic data interpretation and publication guidelines may be referenced from Skou *et al.*[3], Grishaev[36], Jacques and Trewhella[35] and Jacques *et al.*[40].

The basics of SAS: a simple equation with big implications

A very simple relationship links the angular dependence of SAS intensity, I, to the structure of macromolecules in solution, as well as to the bulk properties of a sample. If a sample contains n independent randomly oriented particles, the intensity can be expressed as follows:

$$I(q) = S(q)\sum_i^n [(\Delta\rho_i V_i)^2 P_i(q)] \quad (1)$$

Here $q = 4\pi\sin\theta/\lambda$, where θ is half the scattering angle and λ is the wavelength of incident radiation. This relationship states that the intensity of scattered radiation is the sum of the scattering from each and every individual particle, i, within the illuminated volume of the sample. The angular dependence of $I(q)$ is proportional to several factors, of which the form factor, $P_i(q)$, is perhaps the most interesting factor to the structural biologist. The form factor encodes overall structural information in reciprocal space, which relates to the probable real-space distance distribution between scattering centers within a macromolecule ($p_i(r)$, **Fig. 1**). However, $I(q)$ is also dependent on three other factors: (i) the volume-squared of each particle, V_i^2; (ii) the contrast squared, $\Delta\rho_i^2$, which is the difference in scattering density between the macromolecule and its supporting solvent; and (iii) the structure factor, $S(q)$, which encodes information relating to the correlated motions/distances between particles in solution—i.e., interparticle interactions.

In terms of sample preparation, sample homogeneity, concentration and contrast are the parameters that directly contribute to $I(q)$ and that can be influenced at the laboratory bench. For example, if a sample consists of a mixture of different species in solution—i.e., it is not purified to homogeneity—each species in the mix will have different volumes, contrasts and form factors. As a result, and as equation (1) indicates, the structural parameters extracted from the SAS data will reflect the sum-weighted contribution (not the average) of each species in the mix. Therefore, in order to obtain accurate structural information from macromolecules and to obtain the 3D models of individual proteins, polynucleotides, complexes, assemblies and so on, samples have to be homogeneous and not affected by measurable interparticle interactions (i.e., $S(q) = 1$). If these conditions are met, then the relationship above simplifies to

$$I(q) = N(\Delta\rho V)^2 P(q) \quad (2)$$

where N is the number density of homogeneous particles in the sample. Consequently, under noninteracting (dilute) conditions of a pure sample, the magnitude of $I(q)$ will depend on the particle concentration, volume, contrast, and—importantly—the overall structure and shape. The aim in the wet lab is to optimize sample conditions so that the particle of interest (be it a monomer, dimer, oligomer or complex) is as pure as possible and maintained in a monodisperse state during the course of measurement so that $P(q)$ can be accurately assessed from the scattering intensities. This can be achieved by optimizing concentration, contrast and purity.

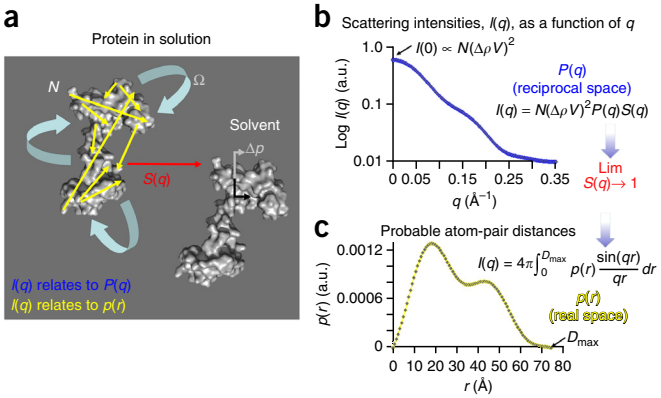


Figure 1 | Scattering basics. (**a**) Macromolecules in solution, e.g., proteins (represented as gray blobs), undergo rotational and translational motion and experience long-range interactions with neighboring particles. The SAS intensities measured from an isotropically tumbling (Ω) monodisperse sample are dependent on a number of factors, of which the form factor, $P(q)$, is of most interest to structural biologists. It is from $P(q)$ that structural parameters and low-resolution models of the macromolecules can be obtained. The form factor of the scattering intensities in reciprocal space relates to the real space distribution, $p(r)$, of all time-preserved, i.e., correlated, pair distances between scattering centers of the molecule (yellow arrows). In the small-angle regime, these correlated distances are otherwise absent in the solvent. However, as all atoms can scatter radiation, solvent-scattering contributions have to be accurately subtracted from the sample scattering to reveal the $P(q)$ of the macromolecules. The magnitude of the intensities will then depend on (i) the number of particles in a sample (N); (ii) the volume squared of the macromolecule (V^2); (iii) the difference in scattering-length density, or the contrast, squared against the solvent ($\Delta\rho^2$); and (iv) scattering arising from correlated distances of closest approach between particles (interparticle interference, or structure factors, $S(q)$). The purity, concentration, contrast and how well a solvent is matched to a sample can be directly controlled during sample preparation. (**b**) SAS data are usually collected on 2D detectors and radially averaged to produce 1D profiles of scattering intensity, $I(q)$, as a function of the magnitude of the scattering vector q. After solvent subtraction, $I(q)$ versus q encodes $P(q)$ from each and every macromolecule in a sample weighted by $N(\Delta\rho V)^2$ and $S(q)$. Longer-distance separations are represented at lower angles and vice versa. At zero angle, $I(0)$, the magnitude of the scattering, is proportionate to the total volume squared and concentration of the macromolecules. (**c**) If $S(q)$ approaches 1—i.e., when the system is infinitely dilute and interparticle effects are absent—modeling the indirect inverse Fourier transform of $I(q)$ versus q produces the real-space $p(r)$ versus r, from which the radius of gyration, R_g, maximum particle dimension, D_{max}, and low-resolution particle shape and structure can be determined. a.u., arbitrary units.

Key points of consideration. The main concepts to keep in mind when preparing samples for macromolecular solution SAS are as follows:

- X-rays are scattered by electrons, whereas neutrons are primarily scattered by atomic nuclei. X-rays are, in general, much more damaging to macromolecules than neutrons, as X-rays can induce chemical changes (e.g., free-radical formation) that can alter the state of a sample over time (e.g., aggregation due to cross-linking).

- All atoms in a sample—not just a macromolecule of interest—have the capacity to absorb or scatter differing amounts of X-rays or neutrons (air, water, sample cells, small chemicals, buffering components and instrument background from slits, windows and so on). As it is impossible to identify where an X-ray or neutron arriving at a detector has scattered from, background scattering intensities have to be subtracted from the sample scattering to reveal the scattering due to the macromolecules. Therefore, at least two measurements must be made under identical conditions: (i) that of the sample (macromolecule + solvent + sample container) and (ii) that of the background (solvent + sample container).
- It is imperative that the solvent in which a macromolecule is suspended is the same as the solvent used to measure the background scattering. If the sample solvent and the background solvent do not match, the resulting subtracted scattering profile will be a mix of scattering intensities derived from both the macromolecules and the mismatched solvent.
- For both X-rays and neutrons, the difference in scattering-length density ($\Delta\rho$) between a macromolecule and the solvent is called the contrast. If $\Delta\rho$ equals zero, then effectively no net coherent scattering will be obtained from a macromolecule after subtracting bulk solvent-scattering contributions (equation (2)), except for weak contributions arising from, for example, internal particle inhomogeneities or the solvation layer around macromolecules. For SAXS, the contrast of a sample depends on the difference between the average electron density of a macromolecule and the average electron density of the aqueous solvent. For SANS, the contrast is the difference between the average neutron-scattering-length density of a macromolecule and the average neutron-scattering-length density of the aqueous solvent. Neutron-scattering lengths are dependent on the isotopic composition of a macromolecule and the solvent.
- For SAXS, the only practical method for altering $\Delta\rho$ is to change the chemical environment of a sample. The X-ray contrast can be altered either by increasing the concentration of small molecules in the solvent or via the addition of electron-dense molecules or heavy atoms to a sample (**Fig. 2**). For SANS, $\Delta\rho$ can be altered by changing the isotopic composition of the sample. The two most abundant isotopes of hydrogen, protium (^{1}H) and deuterium (^{2}H) possess vastly different neutron-scattering lengths. Consequently, $\Delta\rho$ can be manipulated by altering the ^1H_2O:^2H_2O ratio of the supporting solvent or by introducing ^{2}H into recombinant macromolecules at nonexchangeable hydrogen positions (i.e., where ^{2}H is covalently bound to functional groups and not in rapid exchange with the solvent).
- The larger the volume of a particle, the greater the number of correlated distances that exist between scattering centers within the volume of the particle. It is these relatively well-preserved pair-distance correlations, which are otherwise absent in the solvent, that produce SAS intensities at low angles. After background subtraction, the scattering intensity at zero angle, $I(0)$, will represent the sum total scattering from all correlated pair distances weighted by contrast squared. Importantly, for monodisperse systems, $I(0)$ is proportional to the macromolecule volume squared.
- Doubling the concentration of a macromolecule will double the scattering intensity and improve the SAS signal (i.e., the signal-to-noise ratio in the data). However, increasing the concentration too much may lead to correlated distances of closest approach between particles such that $S(q)$, i.e., the structure factor, becomes apparent. Attractive interactions between particles systematically increase structural parameters derived from the experimental data—e.g., the R_g, D_{max} and $I(0)$. Repulsive interactions systematically decrease the structural parameters. Interparticle interference primarily affects data at very low angles, but the contribution can extend well into the useful region of the data, thus complicating interpretation. This is why SAS experiments are usually performed at low solute concentrations, typically below 10 mg ml^{-1} (i.e., one weight percent), and, moreover, why a concentration series needs to be measured in order to extrapolate the data to infinite dilution.
- SAS measurements are performed over a set time period. For X-rays, this could be seconds or milliseconds (synchrotron-SAXS) or minutes to hours (lab-based sources); for neutrons, it is usually minutes to hours. The stability of a sample during the course of data acquisition needs to be ensured.

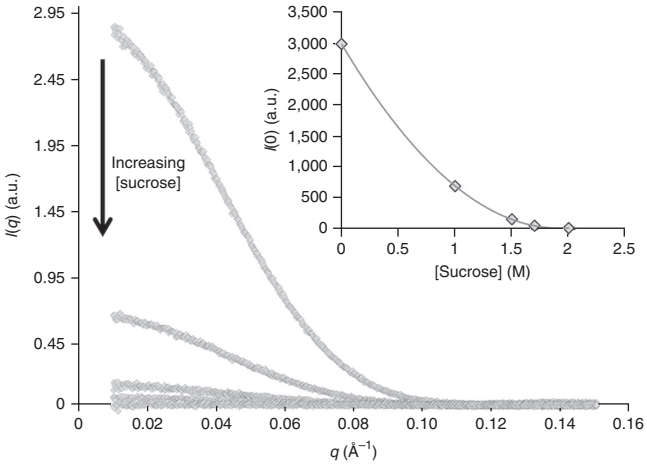


Figure 2 | Decreasing contrast ($\Delta\rho$) and the effect on measured scattering intensities. SAXS data recorded for glucose isomerase with ever-increasing concentrations of sucrose present in the supporting solvent. Increasing the electron density of the solvent relative to the protein results in a large reduction in $I(q)$ caused by a reduction in X-ray contrast, $\Delta\rho$. Inset: the quadratic relationship observed between the calculated total forward scattering at zero angle, $I(0)$, and sucrose concentration. At 2 M sucrose, the protein has effectively been matched out, i.e., $\Delta\rho = 0$. Data were collected at the EMBL P12 BioSAXS beam line[8] of PETRA III, DESY, Hamburg, Germany.

The general points outlined above apply to both SAXS and SANS. However, the physics of SANS—i.e., scattering arising from neutron–nucleus interactions—imposes additional requirements for sample preparation, which are discussed in more detail in the procedures specific to biological SANS experiments.

A quick background for neutron scattering. SANS has the potential to enrich biological structural investigations. Using SANS, the overall low-resolution structure and spatial orientations of macromolecular components of complexes and higher-order assemblies can be extracted from the data—for example, the structure of the ribosome[41], filamentous actin assemblies[42], the subunits of protein–protein complexes[15] and so on. However, compared with SAXS, SANS is experimentally very demanding in terms of sample quantity (typically, tens of microliters for SAXS and hundreds of microliters for SANS), and therefore it is necessary to first answer the question: what is

the specific question SANS can address that other methods, including SAXS, cannot?

The utility of SANS comes from the ability to manipulate the neutron contrast, $\Delta\rho$, of an experimental system without requiring major chemical changes to a sample[38,39,43–45]. Neutron contrast can be adjusted by isotopic substitution, in particular protium–deuterium (^{1}H–^{2}H) substitution, either in the solvent (^1H_2O to ^2H_2O) or via the nonexchangeable (i.e., covalently linked) ^{2}H labeling of a macromolecule. As with SAXS, obtaining homogeneous, monodisperse and pure samples that are not affected by interparticle interactions is also important for SANS. However, there are unique aspects to SANS sample preparation that are influenced by the following:

- The different way neutrons interact with the nucleus of ^{1}H compared with the nuclei of ^{2}H and the other commonly occurring 'biological' isotopes (^{12}C, ^{16}O, ^{14}N, ^{31}P and, mainly, ^{32}S)[46–48].
- The different hydrogen bond strengths of ^{1}H relative to ^{2}H that can alter the solubility of samples or shift the position of disassociation equilibrium of complexes.
- The relatively low flux of neutron sources and the large beam size that require long exposure times and large sample volumes. In comparison with SAXS, radiation damage to a sample is unlikely, but the samples must be time-stable.

SANS basics. All atomic nuclei have a probabilistic capacity to scatter neutrons. The scattering probability, or the scattering cross-section, of a nucleus can be basically pictured as a circle with a radius that relates to what is termed the scattering length of the nucleus. Depending on the nuclear isotope, there can be two scattering cross-sections that describe the neutron–nucleus scattering interaction: coherent and incoherent scattering. Similar to X-rays, the intensities of coherently scattered neutrons relate to the distances between scattering centers within the volume of a particle—i.e., the structure of a macromolecule in solution ($P(q)$). However, incoherently scattered neutrons essentially do not correlate to atom–pair distance separations and therefore scatter radiation independently of q, thus contributing to the measured scattering data as background noise (**Fig. 3**).

^{1}H is unusual in that it has both a negative coherent scattering length as compared with the other major biological isotopes and a very large incoherent scattering length[46–48]. Incoherent scattering provides a structure-uncorrelated background in the form of a constant contribution to all scattering angles, which reduces SANS data quality. Although samples that are rich in ^{1}H will produce intense incoherent background, it is the negative coherent scattering length of ^{1}H that enables the contrast of aqueous biological samples to be altered via ^{1}H–^{2}H isotopic substitution. When perceiving neutrons as waves as opposed to particles, it becomes possible to conceptualize that if two waves of the same wavelength, amplitude and phase add to each other, the result will be a doubling of the wave amplitude. Conversely, if the two waves are 180° out of phase, the waves will cancel each other out. As it happens, the nuclei of deuterium and of the commonly occurring biological isotopes interact with neutrons so that coherently scattered neutrons undergo a phase inversion relative to the phase of the incoming neutron beam[49,50]. This inversion is defined as a positive scattering length (note that for SAXS, the X-ray scattering lengths of all atoms are positive because of the interactions of the charged electrons with the electromagnetic waves).

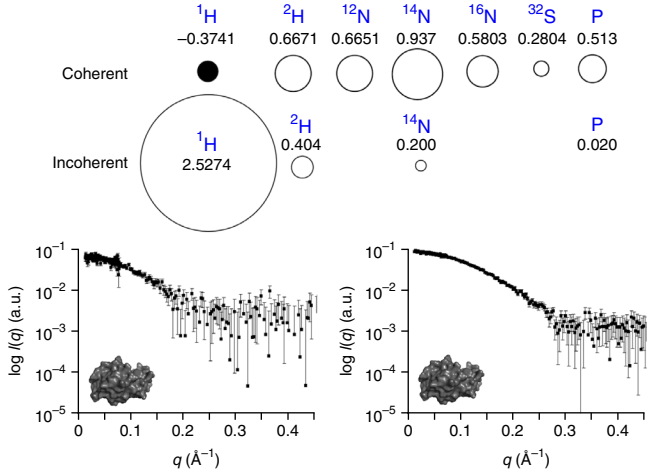


Figure 3 | Coherent and incoherent neutron scattering. Coherent and incoherent neutron cross-sections of the 'biological' elements (σ_c, displayed as circles) and their respective neutron-scattering lengths (b_c, 10^{-12} cm; where $\sigma_c = 4\pi b_c^2$) (ref. 50). ^{1}H has a negative coherent scattering length (represented as a black circle), as compared with deuterium and the other commonly occurring biological isotopes. Coherent scattering arising from correlated distances within a particle's volume produces a scattering profile from which structural information can be extracted. Conversely, incoherent neutron scattering cannot be used to extract shape/structural information and contributes to a SANS profile as 'noise' across all angles. ^{1}H has a considerable incoherent scattering length, the effect of which is demonstrated by the SANS scattering from lysozyme in 100% (vol/vol) ^1H_2O (left), which is considerably noisier than the same sample collected in 100% (vol/vol) ^2H_2O (right). SANS data were collected on the Quokka-SANS instrument at ANSTO[96] using the same neutron wavelength, exposure times, detector distances, instrument geometry, sample path length and protein concentration.

Most isotopes also have positive neutron coherent scattering lengths, but some, e.g., ^{7}Li, ^{48}Ti and ^{55}Mn—and most importantly ^{1}H—do not produce this phase inversion; i.e., the scattering length is negative. As a result, neutrons scattered from ^{1}H are 180° out of phase with scattered neutrons from ^{2}H and the other biological elements. As the neutron contrast in a SANS experiment is simply the difference between the summed coherent scattering lengths per unit volume of a macromolecule compared with that of the solvent—i.e., the difference in average neutron-scattering-length density—and because the scattering length from ^{1}H is negative—the $\Delta\rho$ can be manipulated by simply substituting ^{1}H for ^{2}H in the solvent, macromolecule or both[47].

$\Delta\rho$ = 0: contrast matching. Contrast manipulation increases the information content of an SAS experiment, as scattering contributions from individual components of a complex with different average ^{1}H values per unit volume can be selectively 'matched out' from a scattering profile by altering ratios of ^1H_2O and ^2H_2O in the solvent. Collecting SANS data at a component match point, i.e., at a volume fraction of ^2H_2O that produces $\Delta\rho$ = 0, seems intuitively useless, as the majority of the structural information is effectively removed from a profile. However, if a macromolecule is covalently bound to, or is in complex with, another molecule with a different scattering-length density, then the coherent scattering profile measured at the match point for the first molecule will be derived almost exclusively from the second component. That is, at the match point of macromolecule x it will be possible to obtain

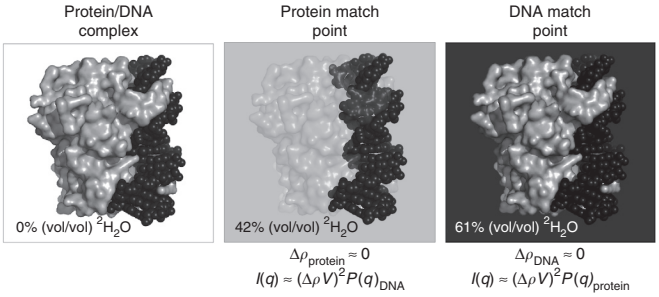


Figure 4 | Principle of contrast matching. If a macromolecular complex consists of individual components that have different average scattering-length densities, it is possible to match out the scattering contributions of a component by placing the complex in a solvent with the same average scattering-length density as that component. Illustrated here is a protein/DNA complex (gray surface and black spheres, respectively). For example, using neutrons, if the complex is placed into ~42% (vol/vol) ^2H_2O, i.e., the protein match point, the measured coherent scattering of the SANS profile will be dominated by the DNA, from which the disposition of the DNA in the complex can be determined. Raising the percentage (vol/vol) of ^2H_2O to 61% matches out the DNA scattering contribution so that the SANS profile is dominated by coherent scattering from the protein.

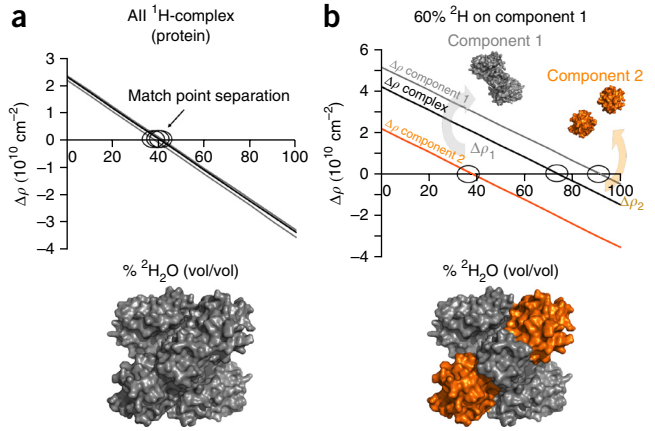


Figure 5 | The effect of nonexchangeable deuterium labeling of a component for SANS with contrast-variation experiments. (**a**) As there is very little difference in the average ^{1}H per unit volume for proteins, the neutron contrasts calculated at different percentages (vol/vol) of ^2H_2O for components comprising a protein–protein complex are almost identical (gray and black linear relations). Consequently, the low-resolution structure restored from a SANS with contrast variation experiment will reflect the shape of the whole complex (gray surface representation). (**b**) Isotopic labeling of a component with nonexchangeable deuterium has a marked effect on the contrast relationships and the separation of match points for the individual components and of the whole complex. In this example, component 1 is labeled on average with 60% nonexchangeable ^{2}H (gray), whereas component 2 remains as a native ^{1}H-protein (orange). When the scattering contributions of the native ^{1}H-protein are matched out (~40% (vol/vol) ^2H_2O; $\Delta\rho_2 = 0$; orange line), the scattering intensities will be derived from ^{2}H component 1, the magnitude of which will change proportionally with each of $\Delta\rho_1^2$ and V_1^2 and $P_1(q)$. On increasing the percentage (vol/vol) of ^2H_2O even further, a point is reached when $\Delta\rho$ for the whole complex approaches zero (~75% (vol/vol); black line), whereby the scattering signal will be exceptionally weak (essentially incoherent scattering and scattering from ^{1}H–^{2}H exchange). Eventually, ^{2}H component 1 will be matched out at a high percentage (vol/vol) of ^2H_2O (~91% (vol/vol) ^2H_2O; $\Delta\rho_1 = 0$; gray line), leaving coherent scattering contributions from the ^{1}H component 2 (proportionally with $\Delta\rho_2^2$ and V_2^2 and $P_2(q)$). From a set of contrast-variation data, it is possible to determine the shapes of the entire complex, the shapes of the individual components and the orientations of the components within the complex.

structural information from macromolecule *y*. Conversely, at the match point of *y*, it will be possible to obtain structural information about *x*. This type of SANS experiment is called contrast matching and is typically performed by choosing the correct ratio of ^1H_2O:^2H_2O in the solvent to match out the components of a complex with different regions of contrast (**Fig. 4**). Different classes of macromolecules have different average isotopic compositions per unit volume—i.e., protein, DNA, carbohydrates and lipids have different ^{1}H:^{12}C:^{16}O:^{14}N:^{31}P: and ^{32}S ratios[48]. Consequently, when focusing on the differences between the average ^{1}H per unit volume of these macromolecules, each class will have a match point at a different percentage (vol/vol) of ^2H_2O in the solvent. Most proteins match out of a SANS profile at between 40 and 45% (vol/vol) ^2H_2O, whereas lipids match out at between 2 and 15% (vol/vol) ^2H_2O and DNA/RNA matches out at ~60 to 70% (vol/vol) ^2H_2O. Many metal nanoparticles, e.g., ferromagnetite, match out at high percentages (vol/vol) of ^2H_2O (e.g., 90–100%), making SANS an attractive option for studying biological macromolecule–metal nanoparticle conjugates. Furthermore, and of particular relevance to this protocol, if a macromolecule is deuterated—i.e., the volume fraction of ^{1}H per unit volume is altered—it becomes possible to control a component's match point (**Fig. 5**).

Contrast variation. Contrast matching can be challenging, as these experiments require the careful formulation of solvents at a specific ^1H_2O:^2H_2O ratio. If SANS data are acquired close to, but not at, the exact match point, the coherent scattering intensities will have contributions from the 'nearly matched' component. For a complex consisting of two components, each with a different contrast in solution (i.e., $\Delta\rho_1$ and $\Delta\rho_2$), equation (2) can be expanded to yield

$$I(q) = N[(\Delta\rho_1 V_1)^2 P_1(q) + (\Delta\rho_2 V_2)^2 P_2(q) + 2\Delta\rho_1 V_1 \Delta\rho_2 V_2 P_{12}(q)] \quad (3)$$

Here it can be seen that $I(q)$ is composed of intensities from the two components, plus an important additional term describing the relationship between them (called the cross-term[48,51]). If a component is not exactly matched, its scattering plus the cross-term will contribute to the observed scattering. SANS with contrast variation experiments overcomes the potential difficulty of exactly matching components and provides additional structural parameters from the cross-term.

SANS with contrast variation data are usually collected from samples using incremental ratios of ^2H_2O in the supporting solvent, often called contrast points, that span the match points of a system. For a two-component complex, there are three match points, $\Delta\rho_1 = 0$, $\Delta\rho_2 = 0$, and for the whole complex, $\Delta\rho_{total} = 0$. At least five, well-spread, contrast points (i.e., scattering curves) are typically measured, preferably above, below and at the individual component match points at different percentages (vol/vol) of ^2H_2O in the solvent. With five such contrast points, there should be sufficient information to extrapolate from the contrast series the form factors of each individual component of the complex, $P_1(q)$ and $P_2(q)$, as well as the cross-term $P_{12}(q)$ that describes the disposition of component 1 relative to component 2. With this information in hand, structural parameters R_g, $I(0)$, $p(r)$ versus r, D_{max} and V, as well as the global structure of the entire complex,

the shapes of the individual components and the spatial orientation between components, can be determined.

Summary

The underlying physics of SAXS and SANS and the relationship between $I(q)$, c, V and $\Delta\rho$ is what ultimately guides sample preparation. Experimenters may not have control over the structure of a macromolecule, but they can control the bulk properties of a sample during its preparation in the laboratory—i.e., sample purity, concentration, monodispersity and contrast. The steps that are necessary to produce quality samples and accurately matched solvent blanks can be challenging. However, the payoffs for optimizing sample conditions can be exceptionally rewarding with respect to improving quality assurance and obtaining additional biophysical information that can reinforce SAS data analysis, modeling and interpretation. This protocol is divided into three main sections:

- Section 1 describes how to assess sample purity and quality for both SAXS and SANS sample preparation.
- Section 2 describes the quantities of material required for SAXS and SANS experiments, including how to estimate sample concentration and MW from SAXS or SANS data.
- Section 3 describes the unique aspects of preparing samples for SANS with contrast matching or SANS with contrast variation experiments.

In addition, **Boxes 1, 2** and **3** detail the practical considerations for performing SEC–SAXS (**Box 1**), how to calculate X-ray and neutron-scattering contrasts (**Box 2**), and the preparation of nonexchangeable ^{2}H-labeled protein for SANS experiments (**Box 3**).

Overview of Section 1

Sample purity and quality. For the sound interpretation of the scattering data, it is vital that the materials undergoing analysis be pure and free of any measurable levels of contamination. As a scattering profile represents the sum of the scattering from each particle in solution (equation (1)), the presence of any contaminants will add to the scattering intensities of a sample. These contaminants will contribute to $I(q)$ at a magnitude that is proportional to the concentration and the volume squared of the contaminant. For example, a sample consisting of a 15-kDa protein purified to 98% that contains 2% of a 100-kDa protein (which does not interact with the 15-kDa protein or itself) will generate a forward-scattering intensity that is almost twice what is expected from a 100% pure 15-kDa sample (**Fig. 6**). Therefore, the presence of high-MW species, including aggregates, aggregates of smaller contaminating particles that coalesce into larger particles, and systems that undergo dynamic nonequilibrium oligomerization or suffer from radiation- or time-induced aggregation, can severely complicate data interpretation and modeling. For successful SAS experiments (especially when developing 3D spatial models that fit the data by shape restoration[18–20] or rigid body modeling[4,21]), it is imperative that macromolecules within a sample be as pure as possible, be monodisperse and remain free of interparticle interference effects. In the PROCEDURE, we first discuss how to assess the sample quality before a SAS experiment (Step 1), and we then discuss how to prepare the sample and buffer for the measurement (Step 2).

Overview of Section 2

Quantity guides for SAXS. One of many considerations when preparing samples for SAXS is the quantity of sample needed for an experiment in order to obtain data with good signal-to-noise ratios and to evaluate the effects of concentration-dependent aggregation or repulsive interparticle interactions. Importantly, obtaining the MW of a sample from the scattering data is a crucial quality assurance step for subsequent data analysis and modeling. As a rule of thumb, sample concentrations for standard synchrotron-based SAXS are usually in the order of 0.1–5.0 mg ml^{-1}. For synchrotron SEC–SAXS, 5–15 mg ml^{-1} (or higher) might be required to compensate for dilution effects through the SEC column that may become important when maintaining the association state of multicomponent complexes (refer to **Box 1**). For far less brilliant laboratory X-ray sources, minimum concentrations for standard measurements are usually 3–10 mg ml^{-1}. The increased concentrations for laboratory-based experiments are necessary to improve the signal-to-noise ratio in the data, especially as particle-scattering intensity decreases markedly with the scattering angle[52]. If the concentration and X-ray flux is too low, those higher q regions of the scattering profile containing information on mid-range atom-pair separations (e.g., domain–domain dispositions, with a resolution of 2 nm and better) can be lost in noise after the buffer subtraction. To determine the concentration of a sample, refer to PROCEDURE Step 7.

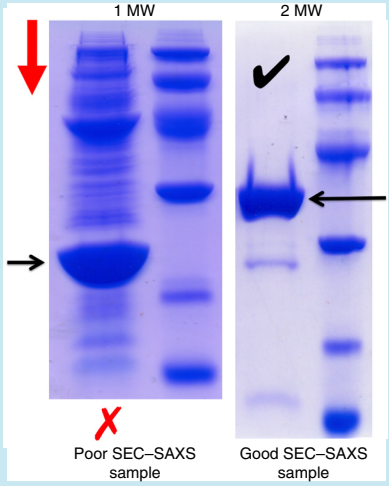
A single SAXS measurement typically requires 5–30 µl of sample (50–100 µl for SEC–SAXS, **Box 1**), with these volumes likely to decrease as microfluidic technologies are introduced[53,54]. However, the volume that can be measured and the volume that is required for reproducible experiments may be quite different. For example, the requirement to match the sample with a solvent before an experiment means that working with tiny volumes can be impractical. Preparing 100–200 µl of sample provides sufficient material to handle it with confidence and provides enough sample to set up a concentration series—e.g., via serial dilution.

- It is always advisable to perform serial dilutions of a sample using the matched solvent as a diluent to assess the effects of $S(q)$ on the scattering intensities (equations (1 and 2)) and to evaluate the association state of multicomponent complexes (whether they are fully formed).

There will always be exceptions to these quantity guides. For example, depending on the macromolecule or complex, it might be possible to analyze a sample at very high concentrations (10–20 mg ml^{-1} and above) as long as $S(q)$ remains insignificant; it may be possible to obtain quality data from a laboratory SAXS source at low concentration (1 mg ml^{-1}) on macromolecules with high MW (e.g., ~300 kDa), or from those that stay stable through extended exposure periods and so on.

Quantity guides for SANS. The quantity of material required for a solution SANS experiment is typically higher than for SAXS. The amount of protein, polynucleotide, complex and so on that needs to be prepared also depends on the type of SANS experiment (e.g., SANS with contrast matching versus SANS with contrast variation; consult the INTRODUCTION and Section 3). In general, 200–500 µl of a sample at 5–10 mg ml^{-1} (plus a corresponding matched buffer) is required for a single 'point' of a SANS contrast series; a full contrast series may require five

Box 1 | In-line SEC–SAXS ● TIMING Buffer preparation + column and detector equilibration, 2–12 h; 1× SEC–SAXS run, 30 min–2 h, depending on the SAXS beam line and SEC column flow rates; data processing, 20 min–2 h.



In-line SEC–SAXS has been successfully integrated as a continuous-flow sample delivery option at a number of synchrotron beam lines, including BioCAT (Advanced Photon Source[98]), SWING (Soleil[99]), the SAXS beam line at the Australian Synchrotron, BM29 at the ESRF and BL23A1 at the NSRRC, Taiwan. At the EMBL P12 beam line (DESY, Hamburg)[8], SEC–SAXS operates in conjunction with a triple detector array that includes RALLS, UV absorption and RI detectors that are placed immediately after the SEC column. The additional detectors are linked in parallel to the SAXS beam line using a mobile phase-flow splitter[93], and they enable the SAXS and independently acquired RI(UV)-RALLS measurements to be directly coupled. By combining the results from laser-light and X-ray scattering with RI or UV measurements, the MW of the separated sample components eluting off the SEC column can be derived.

SEC–SAXS is extremely useful for separating components of already-pure equilibrium systems (e.g., monomer–oligomer interconversion) or removing trace aggregates from a sample immediately before X-ray exposure. However, SEC–SAXS is not a 'cure-all' for every sample and should not be viewed as a purification step but as an analytical procedure to be applied as necessary on a case-by-case basis. For example, the SDS–PAGE evaluation of SEC–SAXS samples in the accompanying gel image shows the results of two samples, 1 and 2. Sample 1, left, contains a target protein of interest (arrow), but too many other protein constituents are present for a successful SEC–SAXS experiment. It would be impossible to use SEC–SAXS to analyze the components of sample 1, as it contains too many contaminants well beyond the resolving power of any SEC column. Sample 2, right, which is almost pure, is a good candidate for SEC-SAXS.

Column resolution (**Table 1**) is determined by the size of the column, the choice of packing matrix, the sample-load volume, the sample flow, solvent conditions and sample purity, all of which need to be evaluated before a SEC–SAXS experiment. If the column resolution is compromised, i.e., the elution peaks 'run into each other', then the SAXS data will also be compromised—i.e., consecutive SAXS data frames collected through the elution will be the sum-weighted contribution from continuously changing ratios of sample component mixtures (equation (1)). However, if the components are well resolved, SEC–SAXS can be invaluable for determining the structure and dispositions of polydisperse systems. The ANTICIPATED RESULTS for a SEC–SAXS experiment are shown in **Figure 14** and refer to SASBDB[17] entries SASDBJ3 and SASDBK3.

The SEC–SAXS method used will depend on the equipment and data processing tools available at a specific beam line. This box provides general advice on how this experiment can be set up and performed.

MATERIALS

- **Protein sample.** SEC–SAXS requires 50–100 µl of protein sample at 5–15 mg ml^{-1}, preferably as pure as possible and filtered through a 0.22- or 0.45-µm spin filter or centrifuged at high speed (10 min; 15–30,000*g*) to remove dust or insoluble aggregates.
- **SEC column.** Refer to **Table 1** regarding column selection.
- **Running buffers.** Make up an excess of buffer to equilibrate the SEC column before and after the SEC–SAXS experiment. Running buffers need to be filtered (using 0.45- or 0.22-µm filters) and degassed. Avoid rapid temperature changes of the column and ensure that the buffer and the column are at the same temperature during the equilibration process. At high-flux SAXS beam lines, it may be necessary to add solution additives—for example, 3–5% (vol/vol) glycerol, 1–2 mM DTT or 1–2 mM ascorbate—to the SEC running buffer to limit radiation damage. Using Tris or HEPES[36,56], instead of phosphate, may also help limit radiation damage (TROUBLESHOOTING; **Fig. 13**).
- **HPLC/FPLC pump flow rate**. Choose a flow rate for the column and equilibrate the column with the running buffer. For SEC–SAXS, flow rates are typically between 0.25 and 0.35 ml min^{-1}. X-ray radiation damage to the sample can occur if the flow rate is too slow. Most commercially available columns have an upper working pressure limit that should not be exceeded.
- **(Optional) Additional detectors**. Where light scattering or spectrophotometric instruments are available, calibrate the concentration (e.g., using RI or UV) and light-scattering intensities (e.g., using RALLS[93] or MALLS) of a MW standard (e.g., for proteins use BSA). The calibrated detectors can then be used to determine the SEC–SAXS sample concentration. The concentration values from UV or RI allow for the processed SAXS data to be placed on a concentration scale for MW determination from $I(0)$ (refer to Section 2, Step 2). If SEC–SAXS UV/RI is combined with MALLS/RALLS, independent estimates of the separated sample components' MW can be obtained that can be used to validate the MW from the SAXS $I(0)$.

PROCEDURE

1. Equilibrate the SEC column, preferably overnight, with SEC running buffer.

▲ CRITICAL STEP The SEC column must be very well equilibrated, typically using 2–8 column volumes of running buffer, before the SEC–SAXS measurement. Extensive column equilibration is required in order to increase the chances of measuring SAXS data corresponding to the matched solvent required for correct background subtraction. Note: a stable UV absorption baseline recorded from the buffer flowing off the SEC column (e.g., at 280 nm) is not an indication that the column has, in fact, equilibrated. For example, a buffer containing 150 mM NaCl will have 280-nm UV absorption properties almost identical to those of a buffer containing 250 mM NaCl, yet these two solutions (that have different electron densities) will produce different SAXS profiles. RI is a more sensitive tool for evaluating whether a column has equilibrated to completion.

(continued)

Box 1 | (Continued)

▲ **CRITICAL STEP** It is strongly advised that SAXS data be collected from a small aliquot of sample (e.g., 10–15 µl) using regular SAXS measurements before SEC–SAXS to assess the radiation susceptibility of the sample. X-ray exposure times for SEC–SAXS may be longer and sample flow rates may be slower, both of which can contribute to increasing the chances of radiation damage. Consequently, if radiation damage is observed using regular SAXS, it is likely that the sample will be damaged during SEC–SAXS (TROUBLESHOOTING; **Fig. 13**).

▲ **CRITICAL STEP** It is necessary to prepare more sample material for SEC–SAXS as compared with regular SAXS, because the sample is diluted 5- to 10-fold as it elutes through the column. To maintain reasonable counting statistics in the SAXS intensities, and to maintain the integrity of macromolecular complexes, high load concentrations are often required to overcome the dilution effects of the column (equation (2), $I(q) \propto N$).

2. Start the SEC–SAXS experiment by injecting the sample onto the column at an appropriate flow rate while at the same time starting the SAXS data collection.

3. (Optional) In parallel with SAXS, begin UV or RI/UV/RALLS or MALLS measurements.

4. Collect SAXS data from the column eluate so that a sufficient number of buffer and sample frames are measured. It is advised to measure SAXS data from the eluting buffer at the beginning and end, and during the SEC–SAXS experiment. Preferably, SAXS data spanning the entire elution profile from the SEC column should be collected.

5. After the sample peak has come off the column, always ensure that the SEC–SAXS experiment runs to completion—i.e., the point at which at least one complete column volume has flowed through the column—or until all sample components have eluted. Flow an additional 0.1–0.25 column volumes of running buffer through the column after the SAXS experiment and before the next sample run. This additional washing ensures that all of the small molecules from the preceding sample are flushed out of the column and do not contaminate the background scattering of the next SEC–SAXS experiment.

6. Assess whether the SAXS sample cell (e.g., sample capillary) is clean after each SEC–SAXS experiment. Compare the (unsubtracted) SAXS profiles measured from the buffer at the very beginning and at the very end of the column elution (e.g., using Correlation Map[93]). If there are differences, clean the SAXS sample cell using a cycle of water–cleaning solution–water. The following are three examples of cleaning solution: (i) 6 M guanidine–HCl, pH 6.5; (ii) 20% (vol/vol) acetic acid; or (iii) 10% (vol/vol) ethanol containing 2% (vol/vol) HellmanexIII.

▲ **CRITICAL STEP** Systematic increases in the scattering intensities of the post SEC–SAXS buffer relative to the initial data frames can indicate that fouling of the sample capillary has occurred. Capillary fouling is often caused by sample components that are susceptible to radiation damage flowing through the X-ray beam; these aggregate and bind to the capillary surface. It is advised to wash the sample capillary between successive SEC–SAXS experiments to reduce the buildup of aggregated material on the internal capillary wall. Aggregate buildup on the capillary makes accurate background subtraction impossible and will contaminate all subsequent SEC–SAXS runs.

7. Select SAXS data frames corresponding to the background scattering for the SEC–SAXS experiment. These frames may be selected from the scattering intensities measured from the solvent/buffer that has flowed through the SEC column. These frames may be—but are not always—close to a sample elution peak.

▲ **CRITICAL STEP** If several data frames are selected and averaged to produce a SAXS profile for the buffer, always ensure that the individual buffer frames are statistically similar before averaging[93]. As samples and buffers run through the SEC column, small-molecule fractionation and/or exchange of the buffer components can occur between the injected sample and the column solvent, as well as between the sample and column matrix—i.e., the beads. Small-molecule fractionation can result in very subtle changes in the SAXS intensities of the buffer as it flows through the column, which may affect the selection of the correct background scattering. To help limit this potential complication (if possible), use dialysis to exchange a sample into SEC running buffer before the SEC–SAXS experiment.

8. Subtract the buffer scattering from each SEC–SAXS data frame. Identify those subtracted frames that correspond to the sample elution peak, for example using *AUTORG*[5] to calculate the R_g and $I(0)$ of the processed data. Make sure that the data have not been over- or undersubtracted (**Fig. 10**), and check that the data frames acquired through an elution peak—after scaling relative to each other (e.g., to concentration)—are statistically similar[93] before any averaging procedure.

9. (Optional) If additional UV or RI detectors have been used to monitor the column elution, correlate the concentration, c mg ml^{-1}, from the detectors to the $I(0)$ from the SAXS and calculate the MW of the eluting components. If (UV)RI-RALLS or MALLS detectors are used, calculate the MW from the light scattering and validate the MW obtained from SAXS $I(0)$. If such detectors are not available, estimate MW (for protein samples) from the particle volume calculated from the SAXS data (refer to Section 2, Step 2).

▲ **CRITICAL STEP** For homogeneous, monodisperse and noninteracting particles, $I(0)/c$, MW and R_g values will be constant. However, obtaining constant values for $I(0)/c$, MW and R_g from SAXS data spanning a SEC elution peak does not always mean that a component is homogeneous and monodisperse. These results depend on the purity of the initial sample and column resolution. BSA, for example, can exist as a mixture that, before SEC, will generate constant $I(0)/c$, MW and R_g values (i.e., using regular SAXS measurements). If this mixture is poorly resolved on a badly prepared or incorrectly chosen SEC column, it is conceivable that the SEC–SAXS data will also produce consistent $I(0)/c$, MW and R_g values through an elution peak. Therefore, before SEC–SAXS, it is advised to perform SEC on a sample to test a selected column's ability to separate the sample components and, if required, alter the solvent conditions (e.g., pH, salt concentration) to optimize separation.

Box 2 | Calculation of X-ray and neutron-scattering contrasts using the *Contrast* module of MULCh ● TIMING 5 min

OVERVIEW OF THE PROCEDURE

MULCh (modules for the analysis of small-angle neutron contrast variation data from biomolecular assemblies[51]) comprises a suite of programs to aid with the analysis of SAS data. The *Contrast* module of MULCh is a specifically tailored for calculating both X-ray and neutron-scattering contrasts of a macromolecular system ($\Delta\rho$). *Contrast* does not require any scattering data as input; it simply uses protein, RNA or DNA sequences in combination with the atomic formulae and concentrations of small molecules in the solvent. Using this information, *Contrast* calculates the X-ray and neutron-scattering-length densities of the macromolecule and solvent (ρ) and subtracts these values to obtain $\Delta\rho$ of the sample.

The contrast values derived from the *Contrast* module can be used to perform the following:

- Assessment of the MW of a macromolecule from $I(0)$ for both SAXS and SANS placed on an absolute scale (cm^{-1}):

$$MW_{sample} = \frac{I(0)_{sample} N_A}{c_{sample}(\Delta\rho \upsilon_{sample})^2}$$

  where c_{sample} is the concentration ($g \cdot cm^{-3}$); υ_{sample} is the partial specific volume of the scattering particle in $cm^3 \cdot g^{-1}$; $\Delta\rho$, the contrast in cm^{-2}; and N_A is Avagadro's number (refer to **Supplementary Method 1**). For proteins and RNA, the partial specific volume, v_{sample}, can be calculated from the primary sequences using the PSV and volume calculator of NucProt[82], http://geometry.molmovdb.org/nucprot/, or it can be obtained from the *Contrast* output. Note that both *Contrast* and NucProt also calculate the volume, V, of macromolecules based on their atomic composition.
- *For SAXS*. Obtain the X-ray-scattering contrast and assess the effect on $\Delta\rho$ when small molecules are added to a solvent. As scattering intensities are proportionate to $\Delta\rho^2$, the addition of high concentrations of small molecules—or the addition of electron-dense molecules—to a sample will reduce the difference in electron density—and thus $\Delta\rho$—between the solvent and a macromolecule of interest. This information may be useful in assessing the effect on the X-ray-scattering intensities (equation (2)) when small molecules that limit radiation damage are added to a sample (e.g., electron-dense polyols, **Fig. 2**).
- *For SANS with contrast matching and contrast variation*. Obtain the neutron-scattering-length density and contrasts of a sample prepared with different percentages (vol/vol) of ^2H_2O concentration in the solvent. From these results, the match points of the sample components can be determined (i.e., the percentage (vol/vol) of ^2H_2O in the solvent that produces $\Delta\rho = 0$), taking into account the percentage of acidic protons that are likely to be in exchange between a macromolecule and the solvent (usually ~90–95%). Note: the value of acidic ^1H–^2H exchange can be altered in the *Contrast* module to evaluate its effect on the sample component match points, which can be useful for deciding on the percentage (vol/vol) of ^2H_2O to use for SANS with contrast matching experiments. In addition, the V and $\Delta\rho$ values from *Contrast* can be used to estimate the change in the overall magnitude of the scattering intensities as components are matched out of the SANS data, equations 2 and 3.
- *For SANS using ^2H-labeled components—pre-production*. If nonexchangeable deuterium labeling of a macromolecule is being considered, *Contrast* can be used to predict the effects of different levels of ^2H labeling on the match-point separation of the components of a sample. Use *Contrast* before setting up a SANS experiment to assess what level of nonexchangeable deuteration is required to obtain the desired sample component match-point separations. These calculations are useful for guiding the production of biodeuterated material before producing ^2H-labeled components (**Box 3**).
- *For SANS using ^2H-labeled components—postproduction*. Calculate the SANS contrasts and the match point of a sample component that has been labeled with deuterium (**Box 3**) using experimentally determined levels of nonexchangeable ^2H from peptide mass fingerprinting results.
- *For SANS*. Calculate the mass density of a $^1H_2O/^2H_2O$ solvent based on the atomic composition. These values can be compared against experimentally determined mass densities from densitometry measurements to check that the percentage (vol/vol) of ^2H_2O of a solvent is correct and to assess experimental SANS contrasts.

The online tool uses a basic copy-and-paste or a simple typing procedure for entering the requisite information. The off-line tool requires a simple text input file (the **Supplementary Data** and **Figure 11** are provided as an example). This simple text input can also be uploaded to the online version of the program. The online and off-line versions of the program generate simple text output files after the calculations that can be reused by either the online or off-line programs. Throughout this procedure, refer to **Figure 11**.

MATERIALS

- A list of solvent/buffer components (atomic formulae) and their concentrations in $mol \cdot l^{-1}$.
- The one-letter amino-acid code or one-letter DNA/RNA code of the macromolecules.
- The atomic formulae of any small molecules bound to the macromolecule of interest—e.g., metal ions, cofactors.

Accessing MULCh. The entire MULCh package, which includes *Contrast*, can be downloaded as an off-line tool (with instructions) or used interactively online via http://smb-research.smb.usyd.edu.au/NCVWeb/.

(continued)

Box 2 | (Continued)

PROCEDURE

1. Enter a project title.
2. Define the solvent. Use the drop-down menu to select the number of dissolved species in the solvent. These are typically small molecules, M. For each small molecule, make sure M is checked, and then enter the atomic formula of each component and its concentration (in mol · l^{-1}) into the appropriate boxes. There is no need to include water, as *Contrast* automatically calculates the change in water concentration as a consequence of adding molecules to aqueous solutions.
3. Define the macromolecule (steps 3–5). Use the drop-down box to select the 'Number of components in subunit 1'. A component can be a protein, DNA, RNA or a small molecule. For example, a protein that consists of only amino acids will have one component. A metalloprotein will have two components—i.e., the protein and the bound metal.
4. Check P = protein, R = RNA, D = DNA or M = molecule, depending on the type of component being described for the subunit. For macromolecules (P, R and D), copy the one-letter code of the entire sequence into the appropriate box. For example, for proteins, check P and then copy the one-letter amino-acid code into the box. For DNA, check D, making sure to copy both the forward- and complementary-strand one-letter sequences. If necessary, define the stoichiometry of the macromolecules in the subunit ($N_{molecules}$). For small molecules that are known to bind to the macromolecule (metal ions, cofactors, and so on), select M, and then type the atomic formula of the molecules, remembering to include the stoichiometry. For example, a protein subunit bound to two calcium ions per monomer is defined as follows: 'Number of components' = 2 (i.e., the macromolecule and calcium), $N_{molecules}$, P (protein), = 1 and $N_{molecules}$, M (calcium), = 2.
5. Define the second subunit of the sample using the 'Number of components in subunit 2'.

▲ **CRITICAL STEP** Samples that are not heterogeneous complexes, for example, lysozyme, glucose isomerase, and so on, are considered by *Contrast* to be a single 'subunit'. In these circumstances, and in order for *Contrast* to complete the $\Delta\rho$ calculation, copy and paste the identical information used to define 'Number of components in subunit 1' into the respective boxes for 'Number of components in subunit 2'. For example, for a tetrameric protein: (i) 'Number of components in subunit 1' = 1; (ii) check P; (iii) list the amino-acid sequence of the monomer; and (iv) set $N_{molecules}$ = 4 (alternatively, input the amino-acid sequence of the tetramer and set $N_{molecules}$ = 1.) Copy the identical information into the 'Number of components in subunit 2' section.

6. (Optional) For SANS using deuterated components, enter the average level of nonexchangeable ^2H incorporated into a macromolecule into the 'Deuteration level' box. Include an estimate of proton–deuterium exchange between the macromolecule and solvent, using the 'fraction of acidic protons accessible to the solvent' box (by default, 0.95).
7. Make sure that the Volume (Å^3) boxes have a number in them, even if it is 0.0. If the volume of a component is known, type in the volume of the component; if the volume is unknown, leave the value at 0.0 and the atomic volume will be calculated automatically.
8. Press submit.

or more points to complete—i.e., 5–25 mg of sample material. However, if a sample remains stable over time and is not adversely affected by multiple rounds of $^1H/^2H$ dialysis exchange (e.g., does not aggregate or change overall shape/structure, oligomerization state and so on), then it is possible to reduce the total quantity of material necessary for a SANS experiment by cycling a sample through solvents with different percentages (vol/vol) of $^1H_2O/^2H_2O$. Use SAXS or dynamic light scattering/static light scattering (DLS/SLS) to check the aggregation state of a sample and, if possible, use SAXS to determine the optimal sample concentration and conditions for a SANS experiment.

Obtaining MW information from the scattering data. Section 2 concludes with how to assess the MW of samples from SAXS and SANS data—i.e., to detect aggregation or significant repulsive interparticle interactions (Step 8).

Overview of Section 3

Preparing samples for SANS experiments. The power of SANS for probing the structures of macromolecules in solution arises from the ability to alter the coherent neutron-scattering contrast of a system without the need to radically alter the chemical environment (which is otherwise necessary for SAXS; refer to the INTRODUCTION). Assessing the neutron contrast for biological samples typically involves ^1H–^2H isotope exchange or substitution in (i) the supporting aqueous solvent, or (ii) (if required) the macromolecule of the sample or (iii) both in combination. As SANS intensities are proportionate to $\Delta\rho^2$ (equations (2 and 3)) and as $\Delta\rho$ can be experimentally controlled by swapping ^1H for ^2H, it becomes possible to isolate the coherent SANS signals produced by the individual components of a multicomponent sample consisting of different regions of neutron-scattering-length density. Consequently, SANS with contrast matching or SANS with contrast variation may be used to determine the low-resolution structure and dispositions of the components of macromolecular complexes and other higher-order assemblies (**Fig. 5**). The following procedure outlines the major practical considerations for designing solution SANS experiments. Steps 9–11 describe how to calculate the contrast match points of macromolecular samples before an experiment—i.e., what percentage (vol/vol) of ^2H_2O is required in a solvent to produce $\Delta\rho = 0$. From here, it can be decided whether deuterium labeling of a component is required to separate the component match points (Steps 12–14). In addition, **Box 3** details the procedure for labeling a protein with nonexchangeable ^2H to alter the coherent neutron-scattering-length density—i.e., the production of biodeuterated material. Steps 15–21 outline aspects of sample quality for SANS with contrast variation experiments, with a particular emphasis on evaluating the solubility of samples in ^2H_2O solvents. Section 3 concludes with how to adjust the pH/pD of $^1H_2O/^2H_2O$-based solvents (Steps 21–24), taking into account the difference in hydrogen bonding strength of ^1H compared with ^2H, and finally Steps 25–34 describe an experimental approach for setting up SANS with contrast matching or contrast variation experiments.

Box 3 | Deuteration of recombinant proteins using a laboratory-based ^2H labeling protocol ● TIMING 5–10 d + additional time if screening of bacterial growth conditions is necessary.

Several SANS facilities offer proposal-based/competitive applications for the production of biodeuterated materials:

- ILL-EMBL Deuteration Laboratory.
  http://www.ill.eu/sites/deuteration/
  http://www.embl.fr/services/deuteration/
- National Deuteration Facility, Australian Nuclear Science and Technology Organisation.
  http://www.ansto.gov.au/ResearchHub/Bragg/Facilities/NationalDeuterationFacility/
- Oak Ridge National Laboratory Bio-Deuteration Facility.
  http://www.csmb.ornl.gov/bdl/

These facilities often use fermenter-based methods to produce large quantities of deuterated bacterial cell pellets containing overexpressed ^2H-labeled recombinant proteins[100]. As ^2H_2O is expensive, and the production of ^2H-labeled components is time-consuming, submitting proposals to these facilities has its obvious benefits. However, it is possible to perform biodeuteration 'in-house' using a simple flask-based procedure in *E. coli* B expression hosts (e.g., *E. coli* Bl21 (DE3)) and using deuterated modified M1 growth medium. A flask-based approach is useful, for example, for obtaining material for assessing the effects of ^2H-labeling on the physical properties of a macromolecule (e.g., stability and solubility in both ^1H_2O and ^2H_2O buffers). The overall procedure can be divided into five stages:

Stage 1: choose the percentage (vol/vol) of ^2H_2O that is required for the final heavy water M1 growth medium to obtain the desired level of nonexchangeable ^2H in the recombinant protein.

▲ CRITICAL STEP The correlation between the percentage (vol/vol) of ^2H_2O used in the M1 growth medium and the incorporation of nonexchangeable ^2H into a protein is not linear. Refer to the table below; this is based on Figure 1 of Leiting *et al.*[86]. For the following example, the desired level of nonexchangeable ^2H for protein X is 60%, which requires a final growth medium of 80% (vol/vol) ^2H_2O using ^1H-glucose as the carbon source.

Percentage (vol/vol) ^2H_2O in M1 medium	Average percentage of nonexchangeable ^2H label in protein
Regular ^1H-glucose carbon source	
0	0
20	12
40	25
60	42
80	61
100	86
Deuterated ^2H glucose carbon source	
80	79
90	88
100	99

Stage 2: adapt the cells to the desired percentage (vol/vol) of ^2H_2O growth medium.

For stage 2, if ampicillin is used as the antibiotic selection agent, it is imperative to resuspend any bacterial cell pellet encountered throughout this procedure in fresh selection medium, as ampicillin slowly degrades over time. Kanamycin is a more culture-stable alternative, as is chloramphenicol. In general, the recommended concentrations of antibiotic for use in ^2H_2O protein expression are as follows: ampicillin, 60–70 μg ml^{-1}; kanamycin, 30–50 μg ml^{-1}; and chloramphenicol, 16–20 μg ml^{-1}.

Stage 3: express the recombinant target in the final ^2H_2O growth medium.

Stage 4: purify the ^2H-labeled recombinant target.

Stage 5: experimentally determine the extent of ^2H labeling in the protein.

The procedure described below starts at stage 2: cell adaption. This step of the procedure spans several days; therefore, the protocol has been divided into several 'mini' procedures that are performed during cell adaption, days 1–5.

With respect to stage 3, i.e., protein expression in ^2H_2O medium, it is assumed that recombinant protein expression has been previously tested using regular bacterial growth medium (e.g., LB medium) and that protein expression requirements, i.e., temperature, inducing agent (if applicable) and antibiotic concentrations or other relevant parameters, such as gene codon optimization, have been screened and optimized for successful recombinant protein overexpression[101]. In addition, if no prior information is available with respect to how well a recombinant protein expresses in ^2H_2O, it is advised to scale down the procedure (from 1 liter) and prepare 50- to 100-ml test cultures. Begin by using the same protein expression parameters as used for 'optimized expression', e.g., in LB medium, to guide the expression in ^2H_2O medium (e.g., temperature, antibiotic concentration and so on). Adjust these parameters if necessary in the test cultures, and then perform the scaled-up procedure using 1 liter of ^2H_2O medium, as described in the text.

(continued)

Box 3 | (Continued)

MATERIALS for stage 2 (cell adaption), day 1

- Plasmid containing the gene of interest and competent *E. coli* B cells
- 1 ml of sterile LB or SOC medium in regular light water
- LB-agar selection plates, made with regular light water supplemented with selection antibiotics

PROCEDURE for stage 2, day 1

1. Transform the desired plasmid into the *E. coli* B cells per the manufacturer's recommendation or using a standard transformation procedure, for example, heat-shock or electroporation, followed by incubation in the growth medium (e.g., 200 µl of LB medium containing no antibiotics) for 1.5–2 h.
2. Plate the transformants out onto the LB-agar selection plates and grow them overnight at 37 °C (or at a predetermined appropriate growth temperature).

MATERIALS for day 2

- 25 ml of sterile-filtered standard LB medium in regular light water, pH 6.5
- 1 × 50-ml sterile Falcon tube (can be purchased as a sterilized product) 1,000× concentrated stocks of your selected antibiotics

PROCEDURE for day 2

1. Pipette 15 ml of sterile LB medium into the 50-ml Falcon tube. Add 15 µl each of the required 1,000× antibiotic solutions. Mix the contents.
2. Remove a 100- to 200-µl aliquot of LB medium and place it in a 1.5-ml Eppendorf tube. Use a sterile loop or pipette tip to scrape 10–15 transformants from the day 1 selection plate into the LB aliquot. Resuspend the cells and use 50 µl of the solution to inoculate the 15 ml of LB medium in the Falcon tube. Close the tube and allow the cells to grow overnight at 37 °C, with shaking.

MATERIALS for day 3

- 10–15 ml of sterile-filtered standard LB medium in regular light water, pH 6.5
- 7.5 ml of sterile-filtered 100% (vol/vol) heavy water, ^2H_2O

PROCEDURE for day 3

1. Centrifuge the cells grown overnight from day 2 (5,000*g* for 10 min), and remove the supernatant. Resuspend the cell pellet in 2–3 ml of fresh 100% LB medium.
2. In a sterile Falcon tube, combine 7.5 ml of ^2H_2O with 7.5 ml of fresh LB medium to produce a 50% (vol/vol) LB/^2H_2O solution. Using the volume graduations printed on the sides of the tubes is adequate to estimate the volume. Add 15 µl each of the required 1,000× stock antibiotic solutions, and mix.
3. Pipette 50 µl of the resuspended cells grown overnight in 100% (vol/vol) LB medium into the 15 ml of 50% (vol/vol) LB/^2H_2O medium. Close the Falcon tube and allow the cells to grow overnight at 37 °C, with shaking.

Cell adaption, day 4

At the beginning of day 4, the cells will have adapted to a 50% (vol/vol) ^2H_2O-LB culture. The process of adapting the cells to a higher percentage (vol/vol) of ^2H_2O begins with the preparation of modified M1 minimal medium (see above) and the ongoing adaption of the cells in small media cultures at ever-increasing percentages (vol/vol) of ^2H_2O (PROCEDURE).

■ **PAUSE POINT** The choice of an appropriate antibiotic concentration in ^2H_2O medium may require additional screening to balance selection versus culture growth time versus culture viability. If the cells have not grown in the 50% (vol/vol) ^2H_2O-LB growth medium from day 3, it may be necessary to repeat the day 2 and 3 procedures using different antibiotic concentrations.

MATERIALS for day 4

- *Inorganic chemical list*: K_2HPO_4, KH_2PO_4, $(NH_4)_2SO_4$, NaCl, $MgCl_2$, $MoNa_2O_4$, $CoCl_2$, $CuSO_4$, $MnCl_2$, $MgSO_4$, $ZnSO_4$, $FeCl_2$, $CaCl_2$, ^2H_2O, ^1H_2O
- *Organic chemical list*: regular ^{1}H-glucose, yeast extract, biotin, thiamine, 1,000× antibiotic stocks
- *Containers list*: Sterile 1 liter and 250-ml Schott bottles (dry), 50-ml sterile Falcon tubes (dry), a sterile smooth-sided and dry conical flask (2.5–3 l) with a stopper

Procedure for preparing minimal medium for use on day 4

▲ **CRITICAL** The steps given below are for the preparation of 1 liter of 1× modified M1 minimal medium in 100% (vol/vol) ^1H_2O and 1 liter of 1× modified M1 minimal medium in 100% (vol/vol) ^2H_2O.

1. *Prepare separate 1H$_2$O and 2H$_2$O phosphate buffers*. Take two dry 1-liter Schott bottles and to each add 10.6 g of K_2HPO_4, 4.94 g of KH_2PO_4, 2 g of $(NH_4)_2SO_4$ and 0.5 g of NaCl. Dissolve the powders in 1 liter of either 100% (vol/vol) ^1H_2O or 100% (vol/vol) ^2H_2O. Adjust the pH of the ^1H_2O phosphate buffer to 6.5; adjust the pH of the ^2H_2O solution to a reading of 6.1 on the pH meter (i.e., pD = 6.5). Use concentrated ^{1}HCl to adjust the pH or pD. If dilute HCL is required for adjusting the pD of the ^2H_2O solution, dilute concentrated ^{1}HCL in ^2H_2O. ^{2}HCl (DCl) can also be purchased for adjusting the pD of the ^2H_2O solution.

For this example, there is no need to use deuterated versions of the potassium or ammonium salts because the total of ^{1}H introduced will not greatly affect the volume fraction of ^2H_2O of the final medium (target = 80%). If perdeuteration is required (i.e., 100% ^{2}H-labeling of a macromolecule), the use of deuterated salts and DCl for pD adjustment is advised. Note: Without a carbon source, the ^1H_2O and ^2H_2O phosphate buffers can be stored at room temperature for several weeks, e.g., after filter-sterilizing the solutions into a sterile, dry Schott bottle. Ensure that the ^2H_2O solution is well sealed to prevent ^1H_2O exchange with the atmosphere.

(continued)

Box 3 | (Continued)

2. *Make a 1,000× stock of vitamin solution in both heavy and light water.* To separate 1-ml volumes of $^{1}H_2O$ and $^{2}H_2O$, add 10 mg of thiamine, 5 mg of biotin and 40 mg of yeast extract. These constituents may not all dissolve in 1 ml. Do not be concerned, and carry on the procedure with the undissolved material. This stock cannot be stored and needs to be used as soon as possible.
3. *Weigh out two separate 4-g amounts of powdered ^{1}H-glucose.* In this example, where the average percentage ^{2}H-labeling of a protein target is 60%, there is no need to use deuterated glucose. For higher percentages of ^{2}H incorporation, a deuterated carbon source may be required in the medium. Refer to the table above and to Leiting *et al.*[86].
4. *Make a 250× stock solution of trace metals in light water.* Dissolve the following in 250 ml of 1 M ^{1}HCl in regular $^{1}H_2O$: 500 mg of $MoNa_2O_4$, 250 mg of $CoCl_2$, 175 mg of $CuSO_4$, 1 g of $MnCl_2$, 8.75 g of $MgSO_4$, 1.25 g of $ZnSO_4$, 1.25 g of $FeCl_2$ and 2.5 g of $CaCl_2$. The solution can be stored for many months at room temperature in a Schott bottle.
5. *Make a 1,000× stock of $MgSO_4$ in heavy water.* Dissolve 2 g of $MgSO_4$ in 5 ml of $^{2}H_2O$ and filter-sterilize the solution.
6. To l liter of the $^{1}H_2O$ and $^{2}H_2O$ phosphate buffers, add 4 g of glucose and 1 ml of the respective ^{1}H or ^{2}H vitamin solutions individually and dissolve. Or, in other words combine the ingredients from Steps 1–3 to make separate 100% (vol/vol) $^{1}H_2O$ and 100% (vol/vol) $^{2}H_2O$ modified M1 minimal media, without adding trace metal or $MgSO_4$ solutions.
7. *Make deuterated medium at the desired percentage (vol/vol) of $^{2}H_2O$.* For this example, the desired percentage (vol/vol) of $^{2}H_2O$ in the final growth medium is 80%. Using a dry measuring cylinder, combine 200 ml of the ^{1}H-medium with 800 ml of the ^{2}H-medium. Filter-sterilize the final 80% (vol/vol) medium and store it in a sterile 1-liter Schott bottle. Do not autoclave and do not adjust the pH or pD of the solution.
8. Make 15-ml 'adaption' cultures at different $^{2}H_2O$ concentrations. Prepare two sterile-filtered 15-ml media solutions at 70% (vol/vol) $^{2}H_2O$ and 80% (vol/vol) $^{2}H_2O$ in 50-ml Falcon tubes using the remaining ^{1}H and ^{2}H-minimal medium.

■ **PAUSE POINT** At this point, both a large 1-liter and two small-scale 15-ml cultures have been prepared. The 15-ml cultures will be used to continue the cell adaption process from day 3 to the finally required 80% (vol/vol) $^{2}H_2O$ environment. The 1 liter of filter-sterilized medium will be used for the protein overexpression experiment on day 6. This solution can be stored at room temperature until needed (but for no longer than ~4 d after the glucose and vitamin solution has been added).

9. (Optional) Make as many small sterile-filtered 80% (vol/vol) $^{2}H_2O$ cultures as possible—e.g., 50 ml in small, stoppered, sterile conical flasks, from any remaining ^{1}H and ^{2}H media. These can be used for test protein expression experiments.

PROCEDURE For the day 4 cell experiment

1. Adapt the cells to 70% (vol/vol) $^{2}H_2O$ minimal medium. Centrifuge the 15-ml cell culture grown overnight in the 50% (vol/vol) $^{2}H_2O$/LB medium from day 3 (5,000*g* for 10 min) and remove the supernatant. Resuspend the cell pellet in 2–3 ml of fresh 70% (vol/vol) $^{2}H_2O$ modified minimal medium.
2. To 15 ml of freshly prepared 70% (vol/vol) $^{2}H_2O$ modified minimal medium in a 50-ml Falcon tube, add 15 µl of the appropriate 1,000× stock solution of antibiotic, as well as 15 µl of the 1,000× $MgSO_4$ stock in $^{2}H_2O$ and 60 µl of the 250× trace metal solution described in Step 4 and 5 of the procedure for preparing minimal medium.

▲ **CRITICAL STEP** Insoluble metal phosphates will form in solution. There is no need for concern; continue the procedure with these insoluble materials in the solution.

3. Add 50–100 µl of the cells adapted to the 50% (vol/vol) $^{2}H_2O$/LB medium to the 70% (vol/vol) $^{2}H_2O$ modified minimal medium and allow it to grow overnight in the closed Falcon tube at 37 °C, with shaking.

PROCEDURE for day 5

Repeat steps 1– 3 of the PROCEDURE from day 4, but this time use 15 ml of 80% (vol/vol) $^{2}H_2O$ modified minimal medium with the antibiotics, $MgSO_4$ and trace metals added to generate 15 ml of an 80% (vol/vol) $^{2}H_2O$-adapted cell culture.

Stage 3: protein expression of the ^{2}H-labeled recombinant target (day 6)

At the beginning of day 6, the bacterial cells should have adapted to growing in 80% (vol/vol) $^{2}H_2O$ modified minimal medium. The 15-ml culture prepared on day 5 will be used to inoculate the 1 liter of 80% (vol/vol) $^{2}H_2O$ medium prepared on day 4. Transfer the filter-sterilized medium to a large, sterile, smooth-sided and dry conical flask (2.5–3 l) in preparation for cell growth and protein expression.

Procedure for day 6

1. Centrifuge the 15-ml culture grown overnight from day 5 in 80% (vol/vol) $^{2}H_2O$ medium at 5,000*g* for 10 min, and resuspend the cell pellet in 1–2 ml of fresh 80% (vol/vol) $^{2}H_2O$ medium.
2. Use the resuspended cells to inoculate the main 1-liter solution of 80% (vol/vol) $^{2}H_2O$ growth medium to an OD_{600nm} value of 0.05– 0.1.
3. Add the 1,000× stock antibiotics (1 ml each), the 250× trace metal (4 ml) and 1,000× $MgSO_4$ solution (1 ml) prepared on day 4. Do not be concerned if precipitates form in the solution; continue with the procedure.
4. Grow the 1-liter cell culture, with orbital shaking, to the mid-log phase of growth, generally between OD_{600nm} values of 0.6 and 0.75.

▲ **CRITICAL STEP** Cell growth in minimal $^{2}H_2O$ medium is very slow compared with that in regular LB medium, and it can take several hours to reach the mid-log phase (e.g., 12 h, as compared with 4 h in LB).

(continued)

Box 3 | (Continued)

5. At the mid-log phase, induce protein expression per the induction method of the plasmid (e.g., the addition of IPTG from 0.1–1 mM) and leave the cultures to express protein for a set time period.
▲ **CRITICAL STEP** As with cell growth to mid-log phase, the expression of the recombinant protein in the ^2H_2O minimal medium may take 2–5 times longer compared with protein expression in regular LB medium or other types of optimized growth conditions (e.g., SOC or Terrific broth). Therefore, use SDS–PAGE to regularly check the level of protein expression during the course of the expression period (e.g., sample 2-, 4-, 6- and 8-h time points, and later, if necessary).
6. Harvest the cells using centrifugation, e.g., 5,000*g* for 15 min. Decant the spent medium from the cell pellet, and transfer the pellet to a storage container (e.g., at –80 °C) or proceed directly to protein purification, stage 4.

Stage 4: protein purification (days 7–9)
It is expected that the purification of the ^2H-labeled protein will follow a scheme similar to that previously determined for purifying the same unlabeled ^1H-protein. Use light-water buffers (there is no need to use ^2H_2O buffers) and purify the ^2H-protein by following the same ^1H-protein purification steps. Adjust the buffer conditions and the protocol, if required, to obtain pure monodisperse protein in solution (main text, Section 1).

Procedure for stage 5 (determination of the average extent of ^2H labeling)
▲ **CRITICAL** Peptide mass fingerprinting is one method used to determine the average level of nonexchangeable ^2H incorporated into the expressed protein target. The experimentally determined value is important for selecting the percentage (vol/vol) of ^2H_2O that should be used in samples for SANS with contrast variation and (especially) contrast matching experiments, to obtain a component match point (i.e., where $\Delta\rho = 0$). We include instructions for how to prepare samples for peptide mass fingerprinting below.
1. Prepare the following materials: SDS–PAGE gel and Tris-glycine-PAGE gel running buffers; 5 µl of unlabeled protein (0.5–1 mg ml^{-1}) in reducing SDS–PAGE loading buffer; 5 µl of ^2H-labeled protein (0.5–1 mg ml^{-1}) in reducing SDS–PAGE loading buffer; and Coomassie blue staining solution and destaining solutions.
2. Perform SDS–PAGE on both the unlabeled and ^2H-labeled proteins.
3. Stain the gel with Comassie blue, and then destain the gel to reveal the protein bands. Wash the destained gel three times in ultrapure (Milli-Q) water for 15 min per wash.
4. Using a scalpel, carefully cut out the bands corresponding to the unlabeled and ^2H-labeled proteins and place the gel fragments into separate Eppendorf tubes.
5. Send the gel slices to a mass spectrometry facility, and request MALDI-TOF peptide mass fingerprinting on both proteins, with mass-fragment (amino-acid sequence) identification and mass analysis.
6. Use the differences between the masses of the peptide fragments obtained from the unlabeled control and those of the ^2H-labeled target to experimentally assess the average level of nonexchangeable ^2H incorporated into the recombinant protein. An example spreadsheet for this calculation is provided as **Supplementary Method 2**.

MATERIALS

REAGENTS

▲ **CRITICAL** The list of the reagents is extensive, and it is assumed that the reader has access to standard laboratory chemicals to make the media, buffers and solutions described below.
- Bacterial growth medium (e.g., lysogeny broth, LB[55]; **Box 3**)
- Buffers for PAGE (e.g., Tris-glycine, Tris-bicine, SDS and so on)
- Buffers for protein purification—e.g., Tris, HEPES and so on
- Buffers for SEC
- Buffers for dialysis
- Proteins: it is assumed that the correct gene of interest has been cloned into an appropriate expression vector (e.g., a plasmid) and that protein overexpression strains of *Escherichia coli* are available (**Box 3**)
- For SANS, access to ^2H_2O is absolutely necessary for contrast matching or variation experiments, noting that ^2H_2O is expensive (~€1,000 per liter)
- Cleaning solutions for SAXS or SANS quartz sample cells (capillaries, cuvettes and so on) include HellmanexIII (Hellma Analytics, cat. no. 9-307-011-4-507), ethanol, guanidine-HCl, acetic acid
- Reducing agents: DTT; β-mercaptoethanol, β ME; or Tris(2-carboxyethyl)phosphine-HCl, TCEP-HCl (e.g., Sigma-Aldrich, cat. no. C4706)
- Additional small molecules for protection against X-ray radiation damage: ascorbic acid and/or glycerol

EQUIPMENT

▲ **CRITICAL** Access to general laboratory equipment and consumables is assumed. Specific equipment for the combined procedures is listed below.
- SDS–PAGE equipment (e.g., from Bio-Rad)
- Dialysis equipment (e.g., SnakeSkin dialysis membrane or Slide-A-Lyzer cassettes)
- Centrifugal spin filters (0.1- to 0.44-µM pore size for filtering out particulates and with nominal MW cutoffs, e.g., 3.5–50 kDa, for protein concentration)
- High-performance liquid chromatography (HPLC, e.g., from Agilent Technologies) system or fast protein liquid chromatography (FPLC, e.g., GE Life Sciences ÄKTA) system
- SEC columns (e.g., from GE Healthcare Life Sciences or Wyatt Technology)
- UV/visible spectrophotometer (e.g., a NanoDrop) or refractometer
- Standalone dynamic and/or static light-scattering instruments, (e.g., from Wyatt Technology or Malvern Instruments) or (optional) in-line SEC multiangle laser light scattering (MALLS)/right-angle laser light scattering (RALLS; e.g., from Wyatt or Malvern) attached to an HPLC or FPLC-SEC system
- Synchrotron bioSAXS beam line or lab-based SAXS instrument (e.g., from Rigaku, Anton Paar, Brucker or Xenocs)
- SANS beam line and SANS quartz sample cells (e.g., water-free QX quartz from Hellma Analytics or SX quartz from Starna) for neutron scattering
- MULCh[51] (can be accessed at http://smb-research.smb.usyd.edu.au/NCVWeb/) or ATSAS[4] (can be downloaded at http://www.embl-hamburg.de/biosaxs/software.html), to perform calculations mentioned in the text
- Additional online tools are mentioned throughout the text

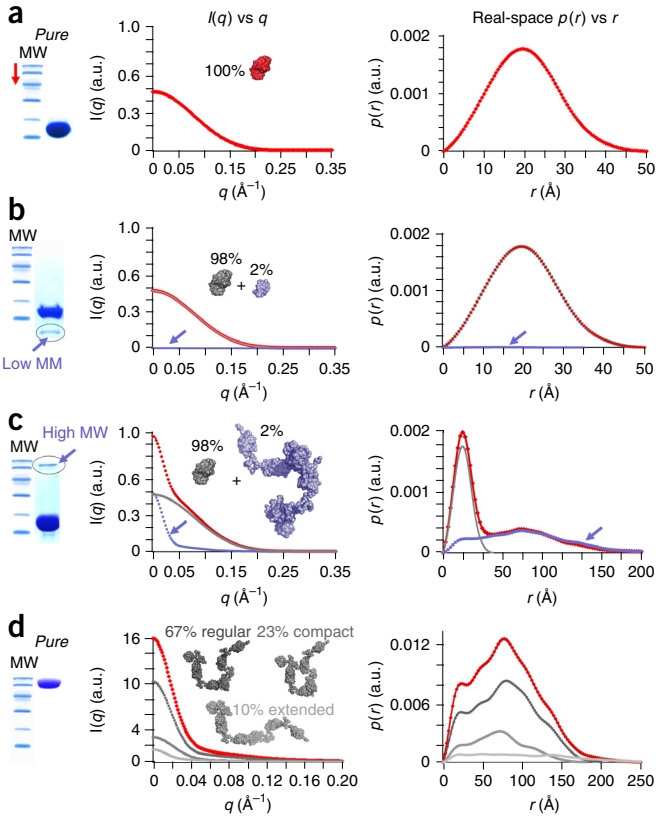


Figure 6 | Sample purity and contaminants (simulated SAXS data and simulated SDS–PAGE). (**a**) The ideal outcome when purifying a sample. Scattering from each individual within a population of pure monodisperse 14-kDa protein sum to produce a total scattering profile (red) from which $p(r)$ versus r can be modeled, which represents the real-space atom-pair distance distribution within a single particle. (**b**) A less ideal situation. If contaminants are present, the total scattering (red) will be composed of the sum of the scattering from each different species in proportion to their volume squared plus concentration. Here, a low-molecular-weight (MW) contaminant (~5 kDa, 2% of the sample, blue) is present in the 14-kDa protein sample (gray). However, the total contribution to the scattering made by the low MW contaminant is small and does not greatly affect $I(q)$ versus q or $p(r)$ versus r. (**c**) Something to avoid. High-MW contaminants have disastrous consequences on $I(q)$ versus q (red). The scattering contributions made by trace ~100-kDa protein (blue) doubles $I(0)$ even though the target 14-kDa protein (gray) is 98% pure. The effect on $p(r)$ versus r is considerable, as it is the sum-weighted contribution made by the 14-kDa protein plus the 100-kDa contaminant. (**d**) A special case: flexibility. A 100-kDa protein is both pure and monomeric. However, the protein is flexible and is composed of three main populations so that the total $P(q)$ determined from the scattering (red) is the sum of the $P(q)$ values from each population (shades of gray). For example, although the extended state comprises only 10% of the total population, the maximum dimension of the measured $p(r)$ versus r (red) will equate to the D_{max} of the most extended state (light gray). Note: the SDS–PAGE gels and scattering profiles used for this figure are for illustrative purposes only and do not represent real data.

PROCEDURE

Section 1. SAS sample purity, quality and preparation of the solvent blank

1| Assess the sample purity and quality before an SAS experiment. The art of biomacromolecular SAS is based in the preparation and characterization of high-quality samples. For SAXS, this includes optimizing conditions that prevent X-ray-induced aggregation[56,57] (refer to the TROUBLESHOOTING section). For SANS, this includes assessing the stability of a sample over the time period required to collect the SANS data. For both SAXS and SANS, the physical aspects of handling samples must also be considered in the context of the preparation, storage and, if required, shipping of samples to distant facilities. For example, unlike X-ray crystallography, in which crystals can be cryoprotected and stored, the simple act of freezing/thawing a sample for SAS, or introducing too many air bubbles, may cause the formation of trace amounts of aggregate that can ruin the interpretation of the scattering data. Therefore, there is a requirement to assess both sample purity and sample stability.

Option A outlines the use of PAGE gels stained with Coommassie blue, which are almost universally used to estimate the ostensive purity of protein samples.

Option B describes the assessment of protein quality by SEC. For precise quantitative analyses of the components present within a sample, the value of SEC cannot be overstated, especially when used in combination with UV spectroscopy and, if possible, with MALLS or RALLS and refractive index (RI) measurements. Most structural biology laboratories have access to SEC-UV equipment as part of standard purification procedures and typically monitor SEC-elution profiles at 280 nm. However, SEC-RALLS-RI or SEC-MALLS-RI instruments (e.g., Wyatt Technology's DAWN HELIOS II plus WyattQELS or Malvern Instruments' Omnisec Reveal and Zetasizer μV) are becoming increasingly useful for the full analytical characterization of sample components (i.e., continuous-flow component separation combined with MW validation and sizing analysis).

Option C details DLS/SLS measurements as techniques for characterizing samples and screening sample conditions. Both stand-alone DLS and SLS (e.g., Wyatt Technology's DynaPro NanoStar or Malvern Instruments' Zetasizer Nano Range) can be used to quickly screen numerous sample environments (e.g., changes in pH and ionic strength) and evaluate sample integrity (e.g., the formation of aggregates). For example, DLS/SLS can be used to assess the effect of adding metal ions, cofactors and so on to a sample, as well as the effects of reducing agents, antioxidants (e.g., sodium ascorbate) or small stabilizing molecules (e.g., 5–10% (vol/vol) glycerol) that may be required to limit the effects of radiation damage in a SAXS experiment (refer to the TROUBLESHOOTING section). The advantage of stand-alone DLS/SLS over SEC-based MALLS-DLS systems is that analyses can be performed using very small sample volumes, and measurements can be completed within minutes.

DLS can also be used to optimize sample conditions; an example procedure is given in option D.

It is important to test different freeze–thaw procedures on the aggregation state of a sample using DLS. In option E, snap-freezing of a sample in liquid N_2 and storage at –80 °C is described; however, similar tests can be performed using samples that undergo snap-freezing on dry ice or slow-freezing and storage at –20 °C (not recommended).

(A) Assessment of sample purity with PAGE

(i) For a SAS sample preparation, use SDS–PAGE to ascertain the presence of contaminants. Undertake further purification steps (e.g., SEC), especially if contaminants have a higher MW than the target of interest (**Fig. 6**). These high-MW contaminants need to be eliminated, as the scattering intensities scale to the volume squared of a macromolecule (equation (1)). Samples purified to 95% without high-MW contaminants present should suffice for most SAS experiments.
▲ **CRITICAL STEP** A single band on an SDS–PAGE gel does not necessarily mean that a sample is monodisperse in solution. Further characterization steps are necessary—e.g., native PAGE and SEC.

(ii) Perform native PAGE (run without SDS) to obtain more information regarding whether a protein sample is predominantly homogeneous or is populated by a range of species (**Fig. 7**).

(iii) For proteins that are often expressed in reducing environments (e.g., internal to a cell), compare SDS and native PAGE with and without reducing agents added to the sample to assess disulfide-mediated oligomerization (e.g., 5 mM DTT; βME or 5–10 mM TCEP-HCl. If necessary, determine the reducing agent concentration (typically 1–10 mM) required to maintain a target of interest in a reduced state (i.e., free of intermolecular cross-links).
▲ **CRITICAL STEP** Make sure that native PAGE gels are cooled—e.g., perform the electrophoresis in a cold room—to prevent heat denaturation of the protein samples.

(B) Assessment of sample quality using SEC

(i) Use SEC-UV to assess the concentration or time-dependent stability of a sample via monitoring the formation of aggregates, higher oligomers and so on.

(ii) Use SEC to obtain information regarding the oligomerization state or the concentration-dependent association between sample components (e.g., of complexes or assemblies). To do this, perform analytical SEC on small aliquots of sample (50–100 µl) through a dilution series using, for example, a Superdex 200 Increase 10/300 GL column (GE Healthcare Life Sciences). Use UV spectrophotometry to monitor changes to the SEC elution profile when changing the load concentration. When choosing the highest sample concentration, make sure that the column does not become overloaded (leading to a loss of separation resolution).

(iii) Evaluate the load concentration required to isolate fully formed (or nearly fully formed) complexes. If performing SEC-UV on a protein–DNA complex, it is important to monitor the UV absorption at two wavelengths, for example at 280 and 260 nm, to demonstrate that a protein–DNA complex has formed and is stable when flowing through the column.

(iv) Optional: Assess the MW and binding stoichiometry of complexes using SEC-UV(RI) combined with MALLS[58] or RALLS. This step can be invaluable when interpreting difficult-to-analyze SAXS samples (as an example, refer to the analysis of the Sda protein from *Bacillus subtilis*[15,59]).

(v) Combine the results obtained from SEC to help interpret the results from PAGE. For example, **Figure 8** shows a protein that, by SDS–PAGE, appears to be pure and monodisperse. However, the SEC elution profile UV trace indicates that the protein is affected by self-aggregation.

Figure 7 | The value of native PAGE. Two proteins, one a wild type (WT) and one a mutant, were analyzed using SDS–PAGE and native PAGE, respectively. The native PAGE result reveals that the mutation radically alters the association state of the protein (from Mokbel *et al.*[97]).

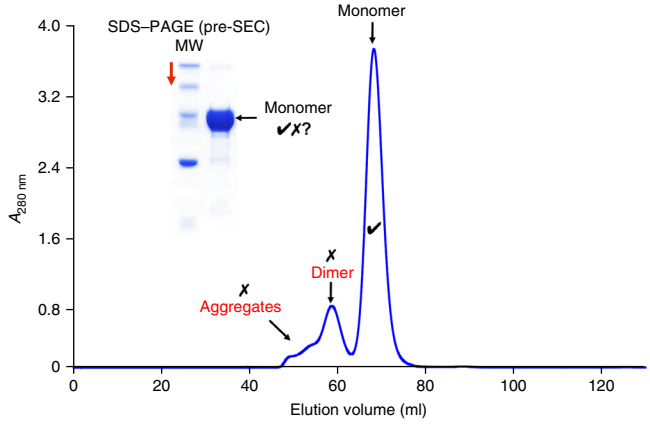


Figure 8 | Sample characterization: SDS–PAGE combined with SEC. The SDS–PAGE result suggests that a protein sample is reasonably pure. However, SDS–PAGE results can be misleading if they are not backed up by further sample characterization. SEC indicates that the sample consists of a heterogeneous population of particles that include self-associated aggregates, dimers and monomers.

(C) Assessment of sample quality using DLS/SLS

(i) Use DLS on minimal quantities (2–10 µl) of material to evaluate the polydispersity and measure the hydrodynamic radius (R_h) of a macromolecule. Use SLS (in combination with accurate concentration estimates) to assess the MW and determine how this may change as a result of altering sample environments. DLS, in particular, is extremely sensitive to the presence of aggregates in a sample that will also negatively affect the results obtained from SAS experiments. The general rule of thumb is that if no aggregates are detected with DLS, then the sample is of sufficiently high quality for SAS.

(ii) Assess the DLS/SLS parameters at various sample concentrations, temperatures and over time to monitor the formation of aggregates. This analysis could prove to be crucial when preparing and isolating monodisperse samples when targets are low-yielding, difficult to produce or expensive.

(iii) Optional: After SAXS measurements have been performed, the R_h obtained from the DLS data can be used to inform subsequent shape analysis. The shape factor R_g/R_h, where R_g is derived from SAXS, offers an additional structural parameter for evaluating the mass distribution of a particle (R_g/R_h of a sphere = 0.78; of flexible random coils (or self-avoiding walks) = 1.44–1.63 (depending on solvent and excluded volume effects[60]); of oblate spheroids = 0.88–0.99; of prolate ellipsoids = 1.36–2.24 (depending on the axial ratio[61]); and of long cylinders or stiff rods = 1.8 to >2 (ref. 61)). In addition, the MW obtained from SLS can be used to validate the MW obtained from SAXS.

(D) Use of DLS to optimize sample conditions (example)

(i) Prepare five 15-µl aliquots of a 2× protein stock solution (e.g., 1–10 mg ml^{-1}) in a buffer of choice.

(ii) To one protein sample, add 15 µl of the same buffer (i.e., without additives) to act as a control. To the remaining four samples, add 15 µl of buffer containing additives at various concentrations (at 1, 2, 4 and 8× the desired final concentration). For example, screen NaCl concentrations from 50 to 300 mM.

(iii) Carefully mix the samples, without introducing air bubbles, and centrifuge at high speed for 5 min (e.g., using a microcentrifuge at 16,000*g*). Carefully remove 5 µl of the sample for the DLS/SLS measurements. For SLS, record the concentration (refer to Section 2).

(iv) Compare the polydispersity and R_h parameters extracted from the DLS measurements for each sample variant. Evaluate any changes in the MW from SLS. Evaluate whether a critical threshold of additive causes aggregates to form in solution.

(v) Store the remaining 10 µl of sample over time (e.g., 3–5 d at 4 °C), and repeat the DLS/SLS measurements to ascertain the time stability and aggregation state with and without additives present.

(E) Use of DLS to determine whether aggregates form on freezing/thawing

(i) Snap-freeze two samples (e.g., 100 µl in Eppendorf tubes) using liquid N_2, and store them indefinitely at –80 °C. Keep aside an aliquot of sample that has not undergone snap-freezing (e.g., store as a liquid at 4 °C).

(ii) Fast-thaw one of the samples (e.g., carefully between your fingertips).

(iii) Slow-thaw the second sample slowly on ice.

(iv) Compare any changes to the sample (e.g., the formation of aggregates) caused by different freeze/thaw procedures against the sample that has not undergone freezing/thawing (**Fig. 9**).

(v) Answer the question: does the sample need to be frozen in the first instance?

(vi) Optional: Use further SEC and PAGE analyses to monitor the effects of different freeze/thaw procedures on the aggregation state of a sample.

Determination of whether samples are affected by concentration or time

2| Determine whether a sample reaches a concentration threshold at which aggregates begin to appear (e.g., with or without freezing/thawing and storage). DLS, PAGE or SEC can be used to perform this analysis. Although increased sample concentration will generate improved signal-to-noise ratios in SAS data (equation (2)), it may be necessary to use lower concentration samples for SAS to avoid the effects of interparticle interactions, especially those that result in the formation of concentration-dependent aggregates.

3| Determine whether the samples are stable through time—i.e., susceptible to aggregation or decomposition into smaller components (e.g., hydrolysis, proteolysis and so on).

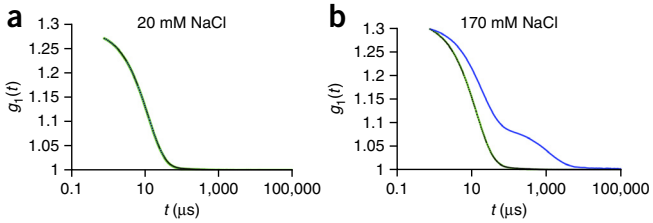


Figure 9 | DLS as a tool for characterizing samples and sample handling. (**a**) DLS autocorrelation functions of lysozyme (9.1 mg ml^{-1}) in low-salt buffer (20 mM NaCl, 40 mM sodium acetate, pH 3.8). A freshly prepared sample (green) is compared with(i) a sample that has undergone snap-freezing in liquid nitrogen and then been quickly thawed (black dots), and (ii) a snap-frozen sample that has been slowly thawed on ice (blue line, obscured). The exponential decay and smooth return to baseline of the autocorrelation data indicate that all three samples are not affected by aggregation, even when stored and handled differently. (**b**) Increasing the NaCl concentration to 170 mM has little effect on the quality of the fresh and quickly thawed samples. However, aggregates are produced if the snap-frozen lysozyme is slowly thawed on ice in the high-salt-concentration buffer (blue). Data were collected using a Wyatt Technology DynaPro NanoStar instrument.

DLS, PAGE or SEC can be used for this analysis. Storage stability for different temperatures and lengths of time may be an important consideration when deciding how to ship samples to distant facilities (e.g., frozen on dry ice or unfrozen on blue ice).

4| Determine whether the macromolecule slowly sticks to the sides of storage tubes. For proteins, this can be monitored by evaluating the concentration of the sample over time in parallel with PAGE (a decrease in concentration or PAGE band intensity is a cause for concern).

5| Assess whether concentration-induced, time-induced or freeze/thaw aggregates and so on can be removed by high-speed centrifugation (e.g., using a microcentrifuge at 16,000*g* or an ultracentrifuge at >30,000*g*), dilution or spin filtration through a centrifugal filter unit (e.g., using 0.1- to 0.45-µm pore-size filter membranes). If not, an additional SEC step may be necessary to remove the contaminating aggregates immediately before an SAS experiment (e.g., using a small Superdex 200 Increase 5/150 GL column, GE Healthcare Life Sciences).

▲ **CRITICAL STEP** The membranes of spin filters used to remove large aggregates or particulate matter can be made from various substrates that include polyethersulfone (PES), modified polyvinylidene fluoride, polytetrafluoroethylene (PTFE), cellulose acetate and cellulose nitrate. If the macromolecules of a sample 'mysteriously disappear' or become increasingly aggregated after spin-filtering, then the membrane is adversely interacting with the sample. If required, test 0.1- to 0.45-µm spin filters made with different substrates.

Obtaining equivalent sample and background solvents

6| For macromolecular solution SAS, it is very important that the scattering contributions made by the background solvent be subtracted from the sample scattering to obtain the scattering from a macromolecule or complex of interest[62]. Inaccuracies in the solvent subtraction will lead to residual solvent terms in the subtracted scattering profile (equation (1)), which can cause perturbations in the structural parameters derived from the data (**Fig. 10**) Therefore, it is essential to produce a solvent blank that is identically matched to the solution of the sample (refer to the INTRODUCTION). In essence, for most scattering experiments, the preparation of the background solvent is nearly as important as preparing the sample. Preparing a matched solvent can be achieved by sample dialysis (option A), by SEC (option B) or by using MW cutoff centrifugal spin filters. In this alternative to dialysis or SEC, the flow-through can be used as an instant sample blank. This approach can work if extreme care is applied.

▲ **CRITICAL STEP** Sample concentration must be considered when choosing the solvent matching method, as the scattering intensities are proportional to the number of homogeneous particles in solution (equation (2)). Dialysis affords more control over the sample concentration, whereas SEC suffers from dilution effects as a sample filters through the column. For high-brilliance synchrotron X-ray sources, sample dilution may not be an issue, but for laboratory-based sources, the overdilution of a sample during SEC may result in compromised signal-to-noise ratios in the data and necessitate extended exposure times (requiring that samples be both radiation- and time-stable). To solve the dilution problem, load concentrated samples onto the SEC column (e.g., 5–15 mg ml^{-1}). However, this is based on the assumption that concentrating a sample does not cause aggregation or result in column overloading, which leads to a loss of SEC resolving power.

▲ **CRITICAL STEP** For SANS experiments, which often require the preparation of several solvents with different percentages (vol/vol) of ^1H_2O:^2H_2O, the buffer exchange method using SEC might not be feasible. Therefore, for SANS experiments, dialysis is advised (refer to Section 3, Step 25).

▲ **CRITICAL STEP** There is a temptation when preparing a solvent blank to simply forgo sample dialysis or SEC and weigh out the components of a new solution that is 'close enough' to the conditions of a sample. This shortcut almost never works. It is difficult to replicate sample solvent conditions (and in particular, density and absorption) for a scattering experiment other than by performing solvent exchange using dialysis or SEC.

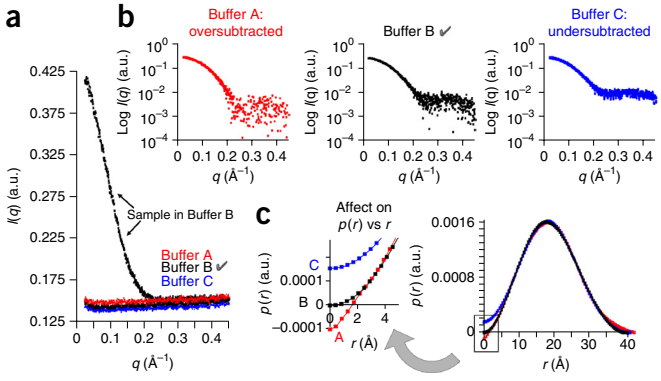


Figure 10 | The importance of obtaining matched sample solvents. (**a**–**c**) SAXS data collected from a pure, monodisperse sample of protein in buffer B and three background solvents, A (red), B (black) and C (blue), before background subtraction. Buffer B was matched with the sample using dialysis, whereas buffer A was derived from a mislabeled tube and buffer C from an old buffer stored in the refrigerator. (**b**) Although the flat solvent-scattering intensities measured for each buffer look similar, only the solvent-matched buffer B allows the correct scattering from the protein to be revealed after solvent subtraction. (**c**) Incorrect buffer subtraction can affect the modeling of $p(r)$ versus r. The correct solvent-matched background (black) causes the profile to intersect at 0,0, whereas under- and oversubtracted buffers result in positive and negative values of $p(r)$ at $r = 0$, respectively. (Note that although real experimental data are presented for the protein, buffer A and buffer B, the data for buffer C have been adjusted for illustrative purposes.)

(A) Perform dialysis to obtain a matched solvent blank

(i) For SAXS—and especially SANS—dialyze a sample against a solvent of choice overnight, making sure that any visible bubbles or air pockets have been removed from the dialysis bag, button or cassette.
▲ **CRITICAL STEP** Using dialysis to perform solvent matching can in some instances be impractical, for example in cases in which a sample is susceptible to slow self-aggregation with time. It may be necessary to test—e.g., using DLS or SEC—that the sample is not affected by time-dependent aggregation during the dialysis procedure.

(ii) Collect scattering data from both the dialyzed sample and the postdialysis buffer. The postdialysis buffer will act as the matched solvent blank for the SAXS or SANS measurements.

(iii) For SAXS, use the postdialysis buffer to dilute the sample to form a concentration series (e.g., 1, 0.75, 0.5 and 0.25×) to assess concentration-dependent interparticle interference effects, $S(q)$, or the disassociation of complexes. Optional: The results derived from SAXS may be used to inform the choice of sample concentration for SANS (refer to Section 3, Steps 17–21. 25 and 26).

(B) Perform SEC to obtain a matched solvent blank

(i) Collect sample fractions corresponding to the separated target of interest eluting from the SEC column.

(ii) Collect aliquots of buffer (e.g., 500 µl) that have passed through the column to act as the matched solvent blank for the SAS experiment. Attempt to collect buffer fractions throughout the SEC run—at the beginning, at the end and as close to the sample elution peak as possible.
▲ **CRITICAL STEP** Use only SEC running buffer that has passed through the SEC column as a matched solvent blank for the SAS experiment, not the buffer from the stock bottle. Limited small-molecule fractionation can occur during SEC, caused by interactions between the solute/solvent/column and the matrix. This fractionation may alter the scattering-length density and absorption properties of a buffer that has run through a column along with the sample, as compared with those of a buffer that has not run through the column.
▲ **CRITICAL STEP** Consider the resolving power of a selected SEC column with respect to sample separation, sample consumption and the selection of the correct solvent for background subtraction. **Table 1** illustrates how column choice is important for resolving components in a sample. Here, BSA, which consists of monomers, dimers and higher oligomers, is separated using three different column matrices. The small Superdex S200 5/150 GL SEC column (GE Healthcare Life Sciences) shows a level of monomer separation (highest peak) from the oligomers, but these monomers are not completely separated from the other self-associated states. Conversely, the Superdex S200 Increase 10/300 GL (GE Healthcare Life Sciences) produces well-resolved peaks, as demonstrated by the UV/RI/RALLS traces returning to a stable baseline between the component elution peaks. The SEC column shows excellent separation of the oligomeric BSA components, whereas the RI/RALLS traces, in particular, are useful for selecting those 'component-free' buffer fractions required for SAS background subtraction.

(iii) Optional: There may be a need to add an expensive, perishable cofactor to a sample (e.g., NADPH) that simply cannot be wasted in the preparation of liters of dialysis or SEC buffer. The best approach in this instance, and one that avoids overly diluting a sample, is to make up a small volume of the additive as a concentrated stock solution (10–50×) in an already-exchanged sample blank. Accurately add a small and equivalent volume (or mass) of the concentrated additive to both the sample and the matched solvent immediately before SAS data collection. If adding an equivalent mass, use a microbalance.

(iv) Optional: Refer to **Box 1**. For SAXS, combine SEC directly with the SAXS measurements—i.e., collect scattering data from the separated sample components immediately after they elute from the SEC column, as well as SAXS data from the elution buffer flowing through the column. Combine the results with additional light-scattering and RI measurements for sample MW validation (ANTICIPATED RESULTS).

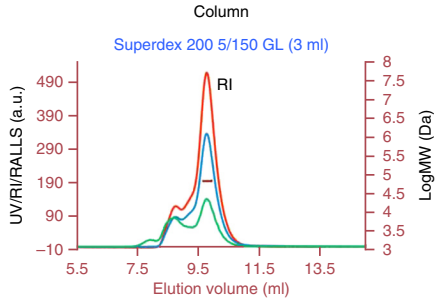
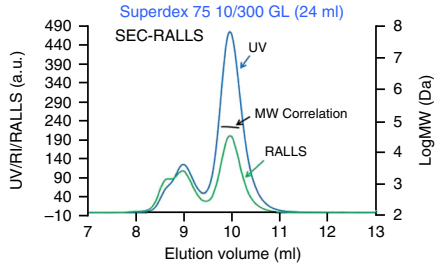
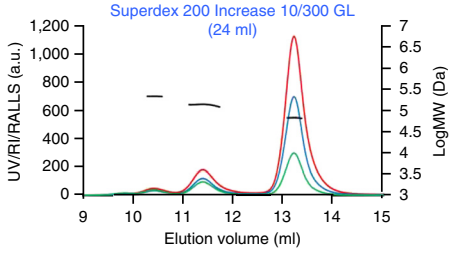
(C) Use of centrifugal spin filters (with extreme care)

▲ **CRITICAL** Spin concentrators can sometimes retain sufficient quantities of small molecules that subtly alter the solvent composition of a sample as compared with the flow-through, resulting in a solvent mismatch. Some filters are manufactured with preservatives coating the membranes (e.g., azide and glycerol) that, if not washed off completely, may introduce unwanted small molecules to a sample. More disastrously, samples can aggregate at the membrane interface as concentration gradients develop during the concentration procedure.

(i) Always choose a new centrifugal spin concentrator with an appropriate MW cutoff. A general rule of thumb is that the minimum MW cutoff should be at least 3–5× less than the MW of the macromolecule of interest. For example, if a protein has a monomer MW of 25 kDa, then use a filter with a MW cutoff of ≤5 kDa.

(ii) Wash the membrane of the filter device carefully with a small aliquot of buffer (e.g., pipetting up and down over the membrane) to remove any small molecules remaining from the manufacturing process. Remove excess buffer and then load the sample.

TABLE 1 | Superdex FPLC column choice and resolving power.

Column	Advantages	Limitations	Example chromatogram
Superdex 200 5/150 GL (3 ml)	• Fast equilibration time and SEC–SAXS experiments (15–20 min) • Low sample consumption (25–40 µl)	• Resolution is too low at the sample concentrations required for SEC–SAXS to separate complex mixtures (e.g., to completely separate monomers, dimers and trimers) • Use as a filter to remove trace aggregates from highly pure samples	
Superdex 75 10/300 GL (24 ml)	• Excellent resolving power for small monomeric proteins (8–50 kDa)	• It is difficult to resolve monomers from oligomers/aggregates, especially if the monomer MW is close to the void volume MW cutoff (i.e., 70 kDa) • Higher sample consumption (50–75 µl) • Longer equilibration time and synchrotron SEC–SAXS experiments (≈ 1–2 h per measurement)	
Superdex 200 Increase 10/300 GL (24 ml)	• Resolves a wide MW range (8–600 kDa). • Excellent separation of complex mixtures, monomers, dimers and trimers • Higher pressures, faster flows. Addition of glycerol to buffers is an option (reduces X-ray damage)	• Higher sample consumption (50–75 µl) • Longer equilibration time and synchrotron SEC-SAXS experiments (≈1 h per measurement)	

UV (blue), RI (red) and RALLS (green) data, and MW correlations (black) were measured using a Malvern Instruments 305 TDA detector/Viscotek FPLC pump system. Columns are from GE Healthcare.

(iii) Centrifuge the sample at a speed or *g* force per the manufacturer's instructions. It is best to concentrate the sample using short multiple spins (e.g., 10 × 2 min) with careful mixing of the sample in between each run as opposed to one long continuous spin (e.g., 20 min). This will help prevent the formation of a concentration gradient and it will reduce the chances of sample aggregation at the membrane/sample interface. Mix the sample between each short spin by carefully pipetting the sample up and down without introducing air bubbles to the solution.

(iv) Once the sample has reached a desired concentration, retrieve the sample and the buffer that has flowed through the membrane. Use the buffer as a solvent blank for the SAXS experiment.

▲ **CRITICAL STEP** If the concentration of the sample in solution decreases or plateaus (i.e., does not increase) during centrifugation, this can be a sign of the protein binding to the filter and, possibly, of the production of irreversible aggregates (e.g., that can be evaluated using DLS). It might be necessary to test different types of membrane substrate to reduce the chances of irreversible binding/sample aggregation. Membranes can be made of regenerated cellulose (e.g., Amicon Ultra, Millipore and Pierce protein concentrators) or PES (e.g., Nanosep and Microsep from PALL Corporation; Vivaspin from GE Healthcare Life Sciences; Corning Spin-X and Pierce PES protein concentrators), as well as modified nylon or hydrophilic polypropylene (Nanosep MF, PALL).

Section 2. Quantity guides, sample concentration and MW

7| *Accurately determine sample concentration.* Aside from acting as a useful tool for monitoring sample-handling procedures, accurate sample concentration measurements are important for the evaluation of SAS data. The determination of

the MW of macromolecules from SAS data is as one of the most important quality assurance steps that links a sample to a scattering profile[35,40,63] and requires the accurate estimation of the sample concentration to within ±10% error. For proteins, amino-acid analysis is perhaps the most accurate way of determining protein concentration, and it provides data to calibrate protein concentration assays, but it can take several days to a week to complete and requires access to skilled personnel and specialized facilities. We therefore recommend either measuring the concentration of a protein spectrophotometrically using absorbency readings at 280 nm or, alternatively, using RI. The RI measurement can, with some adaptation, also be used for the analysis of polynucleotides. The use of conjugating dyes (e.g., Bradford reagent[64]) is generally less accurate for determining protein concentration, except in circumstances in which the dye assays have been confidently standardized (e.g., relative to another technique). Estimating concentration from a known mass of a powdered protein used to reconstitute a solution is generally difficult because of the presence of unknown quantities of salts and other molecules that often accompany powdered protein samples. Assess SAS sample concentrations using option A for polynucleotides and option B for proteins.

(A) Estimation of polynucleotide concentration

(i) As polynucleotides absorb UV light to a great degree at 260 nm, create a dilution series of polynucletotide samples to within the linear response range of a spectrophotometer. For example, double-stranded DNA at a concentration of 1 mg ml^{-1} has an $A_{260\ nm}$ value of ~20 using a 1-cm path length (e.g., perform 10, 20, 40 and 80× dilutions). NanoDrop spectrophotometers (Thermo Fisher Scientific) have shorter path lengths (0.05–1 mm) and can record higher concentrations of polynucleotides without dilution (up to ~15 mg ml^{-1}).

(ii) Measure UV absorbance at $A_{260\ nm}$ for the polynucleotide sample, using the matched solvent to zero the spectrophotometer.

(iii) Divide the $A_{260\ nm}$ absorbance reading by the path length (in cm) and extinction coefficients for double-stranded DNA (~0.020 μg^{-1} ml^{-1} cm^{-1}), single-stranded DNA (~0.027 μg^{-1} ml^{-1} cm^{-1}) or single-stranded RNA (~0.025 μg^{-1} ml^{-1} cm^{-1}), and multiply by the dilution factor to obtain the approximate concentration of the polynucleotide.

(iv) Repeat the measurements at least three times, and obtain the average concentration estimate of the sample.

(B) Estimation of protein concentration

(i) Estimate the protein concentration by measuring the absorbance of a protein sample at 280 nm using the matched solvent to zero the spectrophotometer. Divide the absorbance reading by the path length and the protein extinction coefficient. The extinction coefficient can be calculated from the amino-acid sequence of the protein (e.g., using ProtParam[65]: http://web.expasy.org/protparam/).

(ii) Repeat the measurements at least three times, and determine the average concentration estimate of the sample. ▲ **CRITICAL STEP** Time-dependent or chemical changes in the supporting solvent have to be considered when assessing macromolecular concentration using spectrophotometry. Thiol-reducing agents such as DTT can change their UV absorption characteristics as they undergo oxidation[66] or can interfere directly with dye-based methods. Both DTT and the alternative thiol-reducing agent βME have relatively short half-lives at pH levels >7.5 (~1–20 h depending on temperature[67]), and DTT acts as a chelation agent toward some biologically relevant metal ions such as Zn^{2+} (refs. 68,69), which alters absorption properties and can perturb concentration estimates. TCEP-HCL is a more stable and superior alternative[70] and has negligible effects on $A_{280\ nm}$ readings; however, its effectiveness is compromised in phosphate buffers at neutral pH. Care must be taken to adjust a solvent's pH back to its intended value after TCEP-HCL addition, as it is very acidic.

(iii) Alternatively, perform RI measurements on the protein sample. RI is an extremely useful tool for assessing protein concentration, as the RI increment for a protein (~0.185 ml g^{-1}) is—unlike $A_{280\ nm}$ extinction coefficients—relatively stable against changes in amino-acid sequence composition[71]. The RI increment can also be adjusted for polynucleotides (DNA: ~0.17 ml g^{-1} and RNA: 0.17–0.19 ml g^{-1}). Consequently, RI may be more useful for determining the concentration of, for example, proteins with low $A_{280\ nm}$ extinction coefficients or protein/DNA complexes.

8| *MW analysis.* We have included instructions for determining the MW of macromolecules in solution from SAS data. The MW of a scattering particle can be estimated using a combination of the sample concentration and the extrapolated forward-scattering intensity at zero angle, $I(0)$, derived from the Guinier analysis[72] or from the calculated probable real-space distance distribution, $p(r)$ versus r (refs. 73,74). The procedures outlined below are both concentration-dependent methods: in option A, MW is determined from $I(0)$ using a macromolecule standard (SAXS), and in option B, the MW is determined from $I(0)$ on an absolute scale (SAXS and SANS). For SAXS, use option A to perform concentration-dependent $I(0)$ MW analysis by scaling the sample scattering data to a standard with a known concentration. The standard should have a contrast similar to that of the sample (e.g., use a lysozyme standard for protein samples in aqueous solution[75,76]).

Alternatively, for SAXS—in particular for protein scattering—concentration-independent MW estimates can also be derived from the excluded particle volume, V, which can be computed from the scattering data. Although corrections may be required when assessing MW values based on V for highly extended or flexible particles[77], it is often useful to compare MW values

based on V with concentration-dependent MW estimates based on $I(0)$. Several MW-from-V approaches are available, including the methods of Fischer *et al.*[78] (SAXS-MoW; http://www.if.sc.usp.br/~saxs/), and Rambo and Tainer[79] (through a correlation 'volume') and volume-based MW determinations using *DATPOROD* from the ATSAS software suite[4,5] (where, for proteins, $V_{DATPOROD}/1.6 \sim MW$)). The volume obtained from *ab initio* dummy atom models (DAMs) of proteins that fit the SAXS data (e.g., DAMMIF[6]) can also be used to estimate protein MWs, with the general rule of thumb $MW_{protein} = V_{DAM}/2$. Incorporating MW results from independent light-scattering measurements (SLS/DLS or combined SEC-MALLS or SEC-RALLS) helps one to further validate what is arguably one of the most important overall parameters that may be derived from an SAS investigation (refer to **Box 1** and ANTICIPATED RESULTS).

(A) Determination of MW from $I(0)$, SAXS, using a macromolecule standard

(i) Collect SAXS data from a MW standard and the corresponding solvent blank. Use the same experimental setup, e.g., temperature, exposure time and sample cell (e.g., a capillary). Reduce the scattering data (e.g., radially average 2D data to 1D data) to produce unsubtracted $I(q)$ versus q profiles of the standard and the solvent. Note: for instruments that are not point sources, e.g., Kratky cameras[80], apply the relevant beam geometry corrections to the SAXS data.

(ii) Subtract the solvent scattering from the scattering of the standard to obtain the subtracted 1D $I(q)$ versus q profile of the standard macromolecule in solution.

(iii) Repeat the data collection procedure for a sample with an unknown MW and its corresponding matched solvent. Use the solvent to dilute the sample to form a concentration series (e.g., 8, 6, 4 and 2 mg ml^{-1}). Process the data to obtain the reduced and subtracted $I(q)$ versus q profiles of the sample macromolecule(s) in solution. Apply beam geometry corrections if necessary.

(iv) Calculate $I(0)$ for both the standard and the sample macromolecules using the Guinier approximation (e.g., using *AUTORG*[5]) or from the area under the calculated real-space distance distribution, $p(r)$ versus r (e.g., using *AUTOGNOM*[5]).

(v) Determine the MW of the sample relative to the standard using

$$MW_{sample} = \frac{I(0)_{sample}}{I(0)_{standard}} \times \frac{c_{standard}\Delta\rho^2_{standard}\upsilon^2_{standard}}{c_{sample}\Delta\rho^2_{sample}\upsilon^2_{sample}} \times MW_{standard} \quad (4)$$

where c is the accurately recorded concentration (in wt/vol units) and υ is the partial specific volume (in vol/wt units).

▲ **CRITICAL STEP** If $\Delta\rho_{standard} = \Delta\rho_{sample}$ and the υ values of the standard and sample are similar—which is often the case when standardizing protein SAXS data against a protein standard—the ratio $\Delta\rho^2_{standard}\upsilon^2_{standard}/\Delta\rho^2_{sample}\upsilon^2_{sample}$ in equation (4) is ~1, and therefore it is not necessary to determine the contrasts or partial specific volumes. However, if the standard has a different contrast or partial specific volume as compared with the sample—for example, when comparing a protein standard with a DNA sample, or with a protein sample in glycerol, the ratio will no longer be unity. Under these circumstances, it will be necessary to calculate both $\Delta\rho_{standard}$ and $\Delta\rho_{sample}$ to take into account the differences in electron density of the standard relative to the sample in their corresponding solvents, as well as any differences in υ. Contrast calculations can be performed using the *Contrast* module of MULCh[51] (**Box 2** and **Fig. 11**) or, if an atomic structure is available, by using *CRYSOL*[81]. The partial specific volume can also be estimated using *Contrast*, or, for proteins and RNA, can be calculated using NucProt[82] (http://geometry.molmovdb.org/nucprot/).

(vi) Evaluate whether a systematic decrease or increase in the MW of the sample is observed on changing the sample concentration. An increase in the MW of the sample on increasing concentration is a sign of concentration-dependent oligomerization or aggregation. A significant decrease in the apparent MW of a sample with increasing concentration is typically caused by repulsive interparticle interference.

(vii) Compare the experimentally determined MW from $I(0)$ against the expected MW of a macromolecule, e.g., for proteins, calculated from the amino-acid sequence (using ProtParam[65]). Use the result to evaluate the oligomerization or aggregation state of the sample. Optional: Cross-check the MW result against measurements made from independent light-scattering experiments (e.g., SLS or SEC-MALLS/RI) or, for proteins, the concentration-independent MW based on the estimated particle volume calculated from the SAXS data[4,5,77–79].

▲ **CRITICAL STEP** The standard selected for the MW calibration of SAXS data must be stable in the X-ray beam—i.e., it must not be susceptible to radiation damage (refer to the TROUBLESHOOTING section). X-ray-induced aggregation of the standard will increase the $I(0)_{standard}$ value, resulting in an underestimation of the MW of the sample. In addition, the standard cannot be unduly influenced by repulsive interparticle interference that otherwise decreases the magnitude of the $I(0)_{standard}$ value, resulting in MW overestimates of the sample. If you are unsure, perform SAXS measurements from a concentration series and generate a plot of $I(0)/c$ versus c. The value of $I(0)/c$ should be relatively constant (within error) if $S(q)$ is negligible; a significant positive slope indicates a positive $S(q)$ value, e.g., in the worst-case scenario aggregation; a negative slope suggests a negative $S(q)$ value, i.e., repulsive interactions.

(B) Determination of MW from $I(0)$ for SAXS and SANS, absolute scaling

(i) Perform $I(0)$ MW analysis for SAXS by placing the scattering data of a sample on an absolute scale whereby $I(q)$ has the unit 'per centimeter' (cm^{-1}). For SAXS, absolute scaling is typically performed using the scattering from water as a reference[83].

(ii) Measure SAXS data from pure water in a capillary (or sample cell) and obtain the unsubtracted 1D scattering profile.

(iii) Measure $I(q)$ versus q from the identical, but empty, capillary/sample cell used for the water measurement. Ensure that the empty capillary/sample cell is completely dry and that the X-ray exposure time and temperature are the same as for the water measurement.

(iv) Subtract the empty capillary/sample cell scattering contributions from the water scattering to obtain the subtracted $I(q)$ versus q profile of water alone.

(v) Record the experimental value or 'instrument value' for the forward-scattering intensity of water, $I_{water}(0)_{experimental}$. A simple way to calculate the forward water scattering is to determine the average magnitude of $I(q)$ across a mid-to-high q range—i.e., in the 'flat scattering' region of the water SAXS profile (refer to **Supplementary Method 1**).

(vi) Next, collect SAXS data from the sample and the matched solvent blank in the capillary/sample cell using the same temperature as that used for the water measurement. Subtract the solvent + cell scattering from the sample + cell scattering to obtain the subtracted $I(q)$ vs q profile of the macromolecules of the sample.

(vii) Place the scattering on an absolute scale by multiplying the scattering intensities $I(q)$ of the macromolecules of the sample by the ratio

$$\frac{I_{water}(0)_{standard}}{I_{water}(0)_{experimental}}$$

where $I_{water}(0)_{standard}$ is the known forward X-ray scattering from water at a particular temperature (refer to sheet 2 of **Supplementary Method 1**).

(viii) Determine the $I(0)$ of the macromolecules in the sample from the absolute scaled SAXS data using standard methods—i.e., using the Guinier approximation or from $p(r)$ versus r.

(ix) If $I(0)$ is placed on an absolute scale (cm^{-1}) and c is determined in grams per cubic centimeter, the MW of a macromolecule can be evaluated via

$$MW_{sample} = \frac{I(0)_{sample} N_A}{c_{sample} (\Delta\rho \upsilon_{sample})^2} \quad (5)$$

where N_A is Avogadro's number, and $\Delta\rho\upsilon_{sample}$ is the product of the contrast ($\Delta\rho$, cm^{-2}) and partial specific volume (υ_{sample}, $cm^3 \cdot g^{-1}$) of the macromolecule. Refer to **Supplementary Method 1**, **Box 2**, **Figure 11** and the **Supplementary Data** for further instructions.

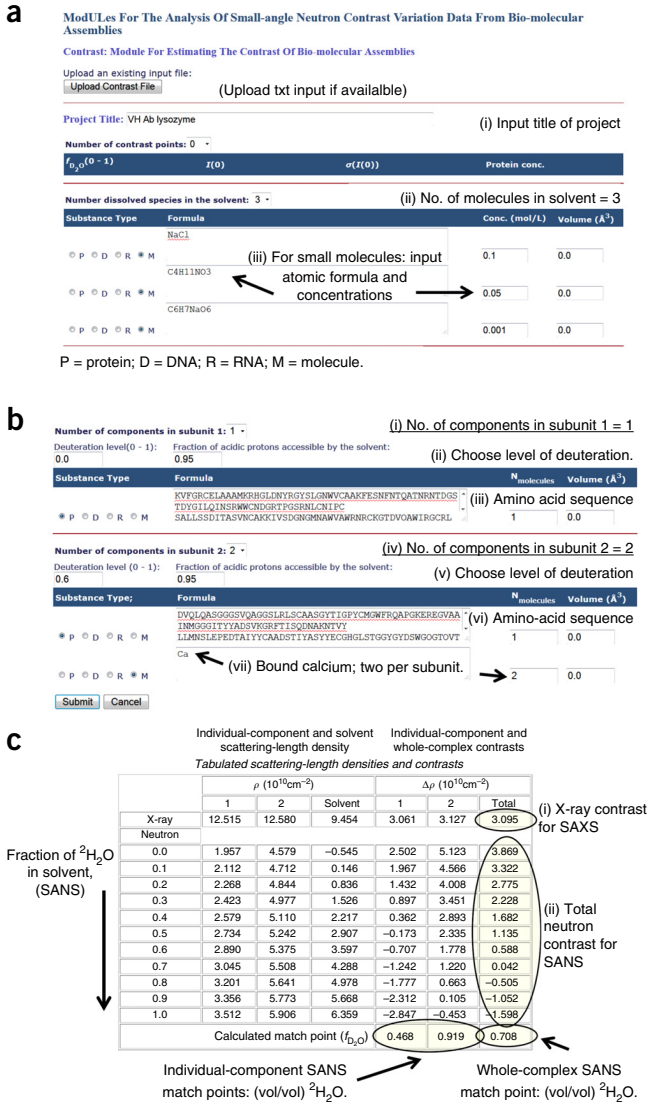


Figure 11 | MULCh calculations of component X-ray and neutron contrasts. The online Web interface of the *Contrast* module of MULCh is shown. (**a**) Defining the solvent. (**b**) Defining the macromolecules of a sample, in this instance, a complex of two proteins, one of which is deuterated and binds two calcium ions. (**c**) The scattering-length and contrast output table from *Contrast* that includes $\Delta\rho$ for SAXS and SANS. For SANS, the magnitude of $\Delta\rho$ is calculated for increasing fractions of ^2H_2O in the solvent. The fraction of ^2H_2O required to obtain the component and whole-complex match points ($\Delta\rho = 0$) are reported.

(x) Perform the absolute scaling of SANS data by normalizing the scattering intensities to the incident beam flux, correcting for sample transmissions and instrument geometry[49,84]. Different SANS facilities will perform the absolute calibration of their instruments using different procedures. Fortunately, most facilities provide their SANS data per centimeter, obviating the need for additional data scaling. Determine the MW from $I(0)$ using the same relationship as in equation (5), substituting $\Delta\rho$ with the coherent neutron-scattering contrasts (calculated by *Contrast*[51]; **Box 2**).

Section 3. Preparation of samples for SANS experiments: calculating the neutron-scattering contrasts of a sample

▲ **CRITICAL** The main issue encountered when using SANS with contrast matching or contrast variation is the separation between the match points of the individual components of a macromolecular complex of a sample relative to the match point of the whole complex, in other words, the points at which the percentage (vol/vol) of ^2H_2O in the solvent produces $\Delta\rho = 0$ for each component and the complex (**Fig. 5**). If the match points of the individual components are too close to that of the whole complex (e.g., within ±10% (vol/vol) ^2H_2O), it may become exceptionally difficult to record sufficiently intense coherent SANS data, especially from samples in ^{1}H-rich solvents that are affected by incoherent ^{1}H scattering (0–50% (vol/vol) ^2H_2O). In such circumstances, match-point separation can be achieved by changing the ^{1}H per unit volume of a component—i.e., by partially or completely deuterating a macromolecule with nonexchangeable ^{2}H that will radically alter the $\Delta\rho$ value (equation (3)). Otherwise-intractable systems become accessible to SANS investigations as a result of deuterium labeling, such as in the analysis of ^{1}H-protein–^{2}H-protein complexes (**Fig. 5**).

9| Use the program *Contrast*, part of the MULCh suite of analytical tools[51], to calculate SANS match points of the individual components and of a whole complex (**Box 2**; **Fig. 11**).

10| The output of *Contrast* includes the calculation of the macromolecular volumes, V, and both SAXS and SANS contrasts, $\Delta\rho$ (cm^2). For SANS, the coherent neutron-scattering contrasts are presented at different fractions of ^2H_2O in the solvent. The volume fraction of ^2H_2O that produces zero contrast is also reported—i.e., the neutron-scattering match points of the components of a complex and of the whole complex.

11| In addition to match-point separation, consider the effect of incoherent scattering from ^{1}H—i.e., the level of 'background noise' that affects the quality of the coherent scattering of a SANS experiment. An increase in incoherent scattering in proportion to the coherent scattering signal, i.e., a worse signal-to-noise ratio, will limit the information content, i.e., the useful angular range of the SANS profiles[52].

The intensity of the coherent scattering signal above the incoherent scattering background is obviously affected by the concentration of ^{1}H in the sample. However, the coherent SANS signal also relates to $(\Delta\rho V)^2$ of a complex and its components. As a component is matched out, the $(\Delta\rho V)^2$ value will limit toward that of the remaining component in the complex (equation (3)). For example, a 50-kDa complex that comprises a 25-kDa ^{1}H-DNA subunit (match point = 70% (vol/vol) ^2H_2O) and a 25-kDa ^{1}H-protein subunit (match point = 42% (vol/vol) ^2H_2O) will likely produce reasonable coherent DNA scattering above the incoherent ^{1}H scattering at the protein match point. However, if the volume ratio of the individual components is more extreme—even if the match points are well separated—for example, a 5-kDa piece of DNA bound to a 45-kDa protein, then the coherent scattering intensities from the small DNA subunit will be very weak at the protein match point and may be 'drowned out' by the incoherent ^{1}H scattering. In general, if a macromolecular component of a complex has 10–15% of the volume (mass), or less, relative to its binding partner(s), it may become challenging to collect quality coherent SANS data from the small component in ^{1}H-rich solvents. The apparent solution to this problem is to increase sample concentration (but this runs the risk of introducing interparticle interference effects) or to increase the neutron exposure time (but this might not be an option given the allocated time on an instrument). The alternative is to isotopically label one of the components with nonexchangeable ^{2}H.

12| If necessary, determine what average level of nonexchangeable ^{2}H labeling is required to obtain SANS match-point separations. This calculation can be achieved by altering the 'deuteration level' parameter in *Contrast* and noting the change in the percentage (vol/vol) of ^2H_2O of the predicted match points from the *Contrast* output (**Box 2**; **Fig. 11**). Alternatively, if an atomic structure is already available, use *CRYSON*[85].

Deuterium labeling will alter the coherent neutron-scattering-length density of a macromolecule and consequently change the magnitude of $\Delta\rho$, producing a shift of the component match point to a different percentage (vol/vol) of ^2H_2O. For example, deuterating a large component of a complex will enable the coherent neutron scattering to be matched out in high-percentage (vol/vol) ^2H_2O solvents (90–100% (vol/vol) ^2H_2O). Under this condition, data can be measured from a small ^{1}H-binding partner in a background with low incoherent scattering. A reversed ^{2}H labeling strategy may also be considered, i.e., the collection of SANS data from a fully deuterated small component—this will increase the magnitude of $\Delta\rho$ in ^{1}H-rich solvents—in complex with a large ^{1}H binding partner. It may be necessary to supplement contrast variation experiments with specialized contrast-matching experiments with alternative ^{2}H labeling strategies to obtain a full set of quality SANS data.

13| If required, isotopically label a macromolecule with deuterium to alter its neutron-scattering-length density. The level and extent of nonexchangeable ^2H incorporated into a macromolecule can be controlled experimentally using biodeuteration[86]. **Box 3** outlines the steps that are necessary for recombinant protein expression in ^2H_2O medium using *E. coli* B bacterial strains—for example, *E. coli* Bl21(DE3). The use of *E. coli* K12 strains (e.g., DH5α) is not recommended.
▲ **CRITICAL STEP** As deuterium labeling can be both time- and labor-intensive, it is advised to initially perform biodeuteration on a small scale (e.g., 50- to 100-ml bacterial cultures) before committing to the production of large quantities of sample. It is necessary to perform feasibility studies to evaluate the levels of recombinant expression, as well as the solubility of the resulting ^2H-labeled product—for example, using SDS–PAGE to test the total and soluble protein content of cell lysates and comparing the results with regular ^1H expression. Optional: Include the results of ^2H-expression trials to support your written proposals to specialized biodeuteration facilities (**Box 3**).

14| After biodeuteration, experimentally determine the average level of nonexchangeable ^2H incorporated into a protein using peptide mass fingerprinting or whole-protein mass spectrometry. The experimental value obtained for the average level of ^2H labeling can be entered into *Contrast* to estimate the expected experimental match points of a complex before a SANS experiment. This will help guide what percentage (vol/vol) of ^2H_2O should be used in solutions to prepare for contrast matching or contrast variation experiments. The basic practical steps for preparing samples for peptide mass fingerprinting are outlined in **Box 3**. **Supplementary Method 2** is provided as an aid for calculating the final average level of nonexchangeable ^2H using the results from peptide mass fingerprinting.

Assessment of sample solubility and stability for SANS experiments
▲ **CRITICAL** What ultimately dictates the success of biological SANS experiments is the stability and solubility of samples in ^2H_2O solutions, as well as the solubility and stability of any ^2H-labeled components. The neutron beam flux, beam size and available time at an instrument also have to be considered, especially in the context of the quantity of material that needs to be prepared for a full SANS with contrast variation series. New sample environments may have to be sought (e.g., altered salt concentration, pH/p^2H and so on) in order to generate conditions in which a complex is stable over time, fully associated, soluble and monodisperse in both ^1H_2O and ^2H_2O solvents.
15| Perform standard sample purity and quality checks—e.g., using SDS–PAGE, native PAGE and SEC—during the purification of the SANS samples. For macromolecules labeled with nonexchangeable ^2H, there is no need to use ^2H_2O buffers during the purification stage. Regular light-water buffers will suffice—i.e., follow the same purification strategy that was used for isolating nonlabeled material.

16| After the sample components have been purified to homogeneity, test the solubility, time stability and effect of storage conditions, with particular emphasis on evaluating the solubility and association state of complexes in ^2H_2O solutions.
▲ **CRITICAL STEP** ^2H_2O generally promotes aggregation. The strength of ^2H-hydrogen bonds is different from that of ^1H-hydrogen bonds, and the solvation layer around a macromolecule has different properties as compared with those of bulk solutions[85,87,88]. The cumulative effects of these differences is that the addition of ^2H_2O to the solvent, or the use of ^2H-labeled components, can affect the solubility, stability and structural dynamics of macromolecular complexes[89–92]. In a worst-case scenario, biodeuterated material might be expressed only in an insoluble form, or components or complexes that are soluble in ^1H_2O might be completely insoluble in ^2H_2O (**Fig. 12**). In addition, it is necessary to test that a complex actually associates in the presence of ^2H. Although the concentration range for SANS (typically, 5–10 mg ml^{-1}) is above the disassociation constant of most physiological complexes, it is prudent to evaluate whether the addition of ^2H_2O to the solvent or the ^2H-labeling of a component (if employed) affects complex formation. Very importantly, a sample must remain soluble, monodisperse and stable across time in both ^1H_2O and ^2H_2O solvents—i.e., it should not aggregate or fall apart during the time required to prepare/store the sample or collect the SANS data.
▲ **CRITICAL STEP** The melting point of pure ^2H_2O is 3.8 °C. Be careful that cold ^2H_2O solutions or samples made in ^2H_2O do not inadvertently freeze when stored in a regular laboratory refrigerator or cold room.

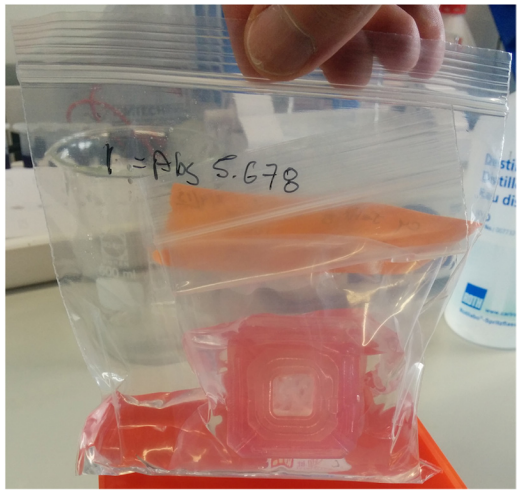


Figure 12 | Dialysis setup for SANS. Two snap-lock or 'sandwich' bags containing 100 ml of 100% (vol/vol) ^2H_2O buffer solution and a 3.5-kDa MW cutoff Slide-A-Lyzer cassette (pink) holding a sample. In this instance, the inner bag leaked and the sample precipitated during dialysis against the high-percentage (vol/vol) ^2H_2O buffer (as shown by the white precipitate).

17| Perform solubility and stability trials of the sample components using material obtained from small-scale purifications before committing to scaled-up procedures. The size of a neutron beam means that 200–500 µl of sample—for each contrast point—may be required—i.e., 5–25 mg of material (or more) for a full-contrast series experiment. Optimization of sample conditions on a small scale is therefore highly recommended.

18| Solubility/stability testing can be achieved by dialyzing test samples (e.g., 50–100 µl) against 0% and 100% (vol/vol) ^2H_2O solvents/buffers (e.g., overnight). Record the concentration of the samples before and after dialysis to assess any significant changes in concentration (e.g., using spectrophotometry; refer to Section 2, Step 7). Note: a significant reduction in concentration of the postdialysis sample may indicate that the sample forms insoluble aggregates that have precipitated out of solution.

19| Use DLS or SLS (as described in Section 1, Step 1, option C and Steps 2–5) to evaluate any significant changes in R_h, MW or polydispersity between the test samples. Use a ^{1}H-macomolecule dialyzed against a 0% (vol/vol) ^2H_2O buffer as a control. If DLS/SLS are not available, SEC in combination with UV spectrophotometry can be used to determine whether soluble aggregates are present in the samples, or if tightly formed complexes disassociate in the presence of ^2H_2O. As ^2H_2O is expensive and will be required for the SEC running buffer, it is recommended to use small 3-ml analytical columns to perform the analysis (e.g., Superdex 200 Increase 5/150 GL column from GE Healthcare Life Sciences).

20| Optional: Perform SAXS on the test samples and their respective solvent blanks, keeping in mind that SAXS is insensitive to ^{1}H–^{2}H isotopic substitution. Evaluate the effects of high-percentage (vol/vol) ^2H_2O solvents on the basic structural parameters of a sample (R_g, D_{max}, $p(r)$ versus r, V and MW). Check that the SAXS scattering profiles of the samples at various ^2H_2O concentrations are similar (e.g., using the Correlation Map[93] method). Changes in the MW of the ^{2}H test samples relative to a ^{1}H control (refer to Section 2, Step 8) are a cause for concern (i.e., aggregation or complex disassociation). Note: an advantage of performing SAXS on the test samples is that the additional dataset can be used in parallel with the SANS results to model the structures of macromolecules.

21| Recommended: Test the integrity of the SANS test samples over a period that reflects the time necessary to acquire SANS data. Neutron flux at SANS beam lines is much lower than that for X-rays, necessitating long exposure periods. Although neutrons are unlikely to cause radiation damage, samples must be time stable during the often extended exposures required to obtain quality data (0.5–4 h per sample and for each matched solvent). For a minimal five-point contrast variation series—i.e., the measurement of five samples and five matched solvent blanks—5–40 h of beam time may be needed to complete an experiment. If it is necessary to improve data quality (especially for samples in which incoherent ^{1}H from the solvent dominates the scattering), quadrupling the SANS collection time should result in an approximate twofold improvement in counting statistics. However, this can greatly extend the length of time samples are exposed in the beam. Therefore, perform a time-course experiment on small aliquots of sample and assess changes in the sample using DLS/DLS, SEC, SAXS and so on, before the SANS experiment.

Adjustment of solvent pH and pD for SANS with contrast matching or contrast variation

▲ **CRITICAL** When making the aqueous solvents for solubility and stability testing, as well as for the final SANS contrast matching or variation experiments, it is important to remember that pH (i.e., for regular 100% (vol/vol) ^{1}H buffers) and p^{2}H (or pD, for 100% (vol/vol) ^2H_2O buffers) are not equivalent[94]. When using glass pH electrodes, the pH and the pD are related via

$$pD = pH_{(\text{measured on pH meter})} + 0.4$$

Consequently, as soon as ^1H_2O and ^2H_2O solutions are mixed to produce solvents with different percentages (vol/vol) of ^2H_2O, it becomes difficult to accurately assess or adjust the pH/pD. The solution to this problem is to make up identical buffers in 100% (vol/vol) ^1H_2O and 100% (vol/vol) ^2H_2O, perform the pH and pD adjustments, taking into account the correction factor, and then mix the 100% (vol/vol) solutions in the appropriate ratios.

22| Obtain two 50-ml (dry) plastic Falcon tubes and individually weigh out the equivalent mass of the components required to make two identical buffers. Be as accurate and precise as possible—for example, use an electronic balance and aim to be within ±10 mg error for each component across both tubes.

▲ **CRITICAL STEP** As all isotopes have the ability to scatter neutrons, the concentrations of ^{12}C, ^{16}O, ^{14}N, ^{31}P, ^{32}S and so on—i.e., the atomic composition—must be identical between the finally made 100% (vol/vol) ^1H_2O and 100% (vol/vol) ^2H_2O buffers, except for the concentrations of ^{1}H and ^{2}H.

23| Dissolve the buffer/solvent components into small volumes (e.g., 10–20 ml) of either 100% (vol/vol) ^1H_2O or 100% (vol/vol) ^2H_2O to generate two tubes of concentrated buffer solution. Transfer the dissolved components from the tubes into an appropriate, and respective, volume of pure ^1H_2O or pure ^2H_2O required to constitute a 1× buffer. Note: bottles or vessels used for 100% (vol/vol) ^2H_2O solutions should always be dry before making ^2H_2O buffers and kept well-sealed to prevent ^1H_2O exchange with the atmosphere.

24| Adjust the pH of the 100% (vol/vol) ^1H_2O buffer and then adjust the pD of the 100% (vol/vol) ^2H_2O buffer to the desired value, taking into account the correction factor of 0.4. For example, when adjusting the pD of a 100% (vol/vol) ^2H_2O buffer to 7.0, the pH meter should read 6.6. Preferably, use ^2HCl or NaO^2H (often called DCl and NaOD) to adjust the pD.

25| To make solvents with different percentages (vol/vol) of ^2H_2O, mix the pH-adjusted 100% (vol/vol) ^1H_2O buffer and the pD-adjusted 100% (vol/vol) ^2H_2O buffer to the appropriate ratio without any further pH/pD adjustments.
▲ **CRITICAL STEP** If it is necessary to add a cofactor, expensive reagent or a component at low concentration, it is advised to accurately weigh out identical quantities of material and make up 10× or 20× stocks (e.g., in 1 ml) for both 100% (vol/vol) ^1H_2O and 100% (vol/vol) ^2H_2O, respectively. The appropriate volume of the concentrated stock can then be added to the main 1× 100% (vol/vol) ^1H_2O or 100% (vol/vol) ^2H_2O buffer solutions before pH/pD adjustment.

Setup of a SANS with contrast variation series

26| The simplest way to set up a SANS contrast series is via the dialysis of the samples against a solution containing different percentages (vol/vol) of ^2H_2O. Dialysis is typically performed overnight to complete ^1H–^2H exchange between the solvent and the sample. There are two main options available for setting up the samples for a contrast experiment, depending on whether the amount of sample material is limited (0.6–1 ml; option A) or plentiful (e.g., 2–2.5 ml; option B).

The key things to remember when setting up the dialysis are that it is important to be consistent when handling the sample, and to be consistent and accurate when making up the dialysis solutions, and that it is necessary to eliminate (to the extent possible) contaminating or unknown quantities of ^1H that would otherwise alter the contrast and/or introduce incoherent scattering noise in a SANS profile.

For both options A and B, prepare a master stock of a macromolecule, complex or assembly in regular ^1H_2O buffer at the concentration selected for the SANS experiment (e.g., 7 mg ml^{-1}). Divide the master stock into smaller subsamples for the subsequent dialysis and mixing steps. As ^1H–^2H exchange can occur across different time scales, overnight dialysis is recommended (or a minimum of 8 h).
▲ **CRITICAL STEP** If a deuterated component is being used for the SANS experiment, it should be derived from the same batch of biodeuterated material (i.e., the same batch of cells; **Box 3**). Avoid making multiple small samples derived from different protein preparations that can be affected by differences in concentration, mixing errors and deuteration levels.

(A) Preparation of limited sample material, e.g., for contrast matching

(i) Separately dialyze two aliquots of an identical sample (e.g., 300–500 µl) against 100 ml of 0% (vol/vol) ^2H_2O and 100 ml of 100% (vol/vol) ^2H_2O solutions, respectively.

(ii) After the dialysis is completed, use an accurate pipette to carefully mix the postdialysis samples together in the appropriate volume proportions necessary to obtain the desired percentage (vol/vol) of ^2H_2O in the sample. It is also necessary to carefully mix the postdialysis 0% and 100% (vol/vol) ^2H_2O buffers using the same volume ratios as the samples to act as the matched solvent blanks for the SANS measurements.

(B) Preparation of plentiful sample material, e.g., for contrast variation

(i) From 2–2.5 ml of a master stock, separately dialyze 200- to 500-µl aliquots of sample overnight against 100 ml of each respective percentage (vol/vol) of ^2H_2O solvent required for the contrast series. This typically includes preparing ^2H_2O solutions above, below and at the two component match points (e.g., 0%, 20% 42% 68%, 90% and 100% (vol/vol) ^2H_2O). Although this approach uses more material than option A, it avoids bad pipetting or mixing errors, reduces the chances of formulating mismatched solvents and prevents potential ^2H_2O 'mixing-shock' that could destabilize/subtly aggregate a sample.

(ii) Centrifuge the dialyzed samples at high speed (30,000*g*, 10 min) to remove any insoluble material.

(iii) Use 200–500 µl of the postdialysis buffers as the matched solvent blanks for each percentage (vol/vol) ^2H_2O contrast point, making sure to either centrifuge the buffer at high speed (30,000*g*) or filter it through a 0.22-µm filter to remove particulates.

27| *Assess sample concentration*. After dialysis is complete, record the concentration of each sample to assess any changes caused by the dialysis procedure. The concentration values will be required for determining the MW of the samples from $I(0)$ at each SANS contrast point (equation 5; **Box 2**).
▲ **CRITICAL STEP** Wherever possible, all consumable materials (volumetric pipettes for aliquotting buffers, pipette tips, dialysis cassettes, loading syringes, needles, Eppendorf tubes and so on) should be dry and dedicated to the separate

handling of ^1H_2O and ^2H_2O samples and buffers. This reduces the chances of ^1H–^2H cross-contamination. Small-volume dialysis cassettes (500 µl) with low-MW cutoffs, e.g., 3.5 kDa, are sufficiently strong to be loaded dry (e.g., Slide-A-Lyzer Dialysis Cassettes from Thermo Fisher Scientific, cat. no. 66333.) If dialysis cassettes are not available, hydrate dialysis membranes in each individual percentage (vol/vol) ^2H_2O solution. Take care not to cross-contaminate the membranes with different ^1H-containing solutions. Remove any excess solution from the membranes (e.g., using a Kimwipe) before loading the sample, to prevent sample dilution.

28| Eliminate unknown quantities of ^1H during dialysis and data collection. Perform sample dialysis in airtight resealable plastic snap-lock/sandwich bags. This approach reduces the slow contamination of the samples from unknown quantities of atmospheric water vapor while using minimal amounts of buffer to cover the dialysis cassette (**Fig. 12**). The only ^1H introduced into the contrast series will be the consistent amount derived from the sample master stock.

29| Decide which SANS sample cells will be best for your experiment.
▲ **CRITICAL STEP** SANS samples are typically irradiated with a large beam (8–15 mm across) using 18- to 20-mm round 'banjo' or 12 × 45-mm rectangular-shaped quartz cells with a typical path length of 1 mm (e.g., 1/SX/1 from Starna or 100-QX-1 from Hellma). Cells with a 2-mm path length can be useful for collecting SANS data from low-concentration samples, but they are reserved for systems containing higher proportions of ^2H_2O that have low incoherent scattering and lower absorption of the specimen. The disadvantage of having such large sample cells is the obvious requirement for large sample volumes. Less obvious is that the cells are prone to inadvertent ^1H contamination.

30| Check that the sample cells are completely dry.
▲ **CRITICAL STEP** The SANS sample cells must be loaded clean and dry (using a dry gas stream or oven) so that any residual ^1H_2O from a previous wash does not carry over to the next measurement.

31| Equilibrate the samples to the temperature at which the SANS experiment will be performed.
▲ **CRITICAL STEP** The large external surface area of the SANS sample cells can condense ^1H_2O vapor from the atmosphere, i.e., can 'fog-up', if cold samples are moved into a warm humid environment. It is thus important to equilibrate the samples and buffers to the temperature of the SANS experiment before loading the cells into the instrument. This will reduce the chances of external fogging, as well as the slow formation of bubbles on the internal surface of the quartz derived from dissolved gasses in the solvent. Both ^1H contamination from fog and bubble formation will introduce unwanted scattering and alter the contrast.

32| Load the samples for each contrast point and the respective postdialysis buffers into SANS sample cells using a pipette with a plastic tip (e.g., a Gilson P200 Pipetman). Gel-loading tips with a thin tapered end can aid loading. Never use a metal needle.
▲ **CRITICAL STEP** Avoid introducing air bubbles when loading the SANS sample cells. Small bubbles can be removed by gently tapping the loaded cell on a hard surface. Additional sample degassing using low-powered sonication can also be used to remove bubbles; however, caution must be applied when using sonication: flawed or scratched sample cells can shatter. Alternatively, place the loaded SANS sample cell into a 50-ml Falcon tube that is packed at the end with a dry Kimwipe or tissue and spin at low speed (less than 500*g*) for 1 min.

33| Seal the samples with Parafilm or a similar material.
▲ **CRITICAL STEP** A full SANS with contrast variation series may take 5–40 h to complete. Therefore, it is important to seal the sample cells during the course of measurement to prevent evaporation and the exchange of ^1H_2O in the atmosphere. Parafilm, or a combination of Parafilm plus thread sealing tape (PTFE), can be used to seal the SANS sample cells.

34| Remove any fingerprints or other residues from the external surface of the cells using a Kimwipe before loading in the SANS instrument.

35| After, or during, the SANS measurement—if possible—perform mass densitometry measurements of each solvent blank of the contrast series. Inject 1–2 ml of the postdialysis buffers that have not been exposed to the neutron beam into a density meter (e.g., an Anton Paar DMA 500 Density Meter). These data can be used to assess the experimental percentage (vol/vol) of ^2H_2O of each contrast point and can be correlated against the neutron beam transmissions. Alternatively, the neutron beam transmissions themselves can be used to evaluate the deuterium content of the solvent to obtain experimental estimates of the percentage (vol/vol) of ^2H_2O (ref. 95). These solvent density and transmission measurements can be compared with the mass density versus $\Delta\rho$ function, calculated using the *Contrast* module of MULCh[51], to obtain estimates of the experimental neutron contrasts, which are useful for subsequent data analysis and modeling.

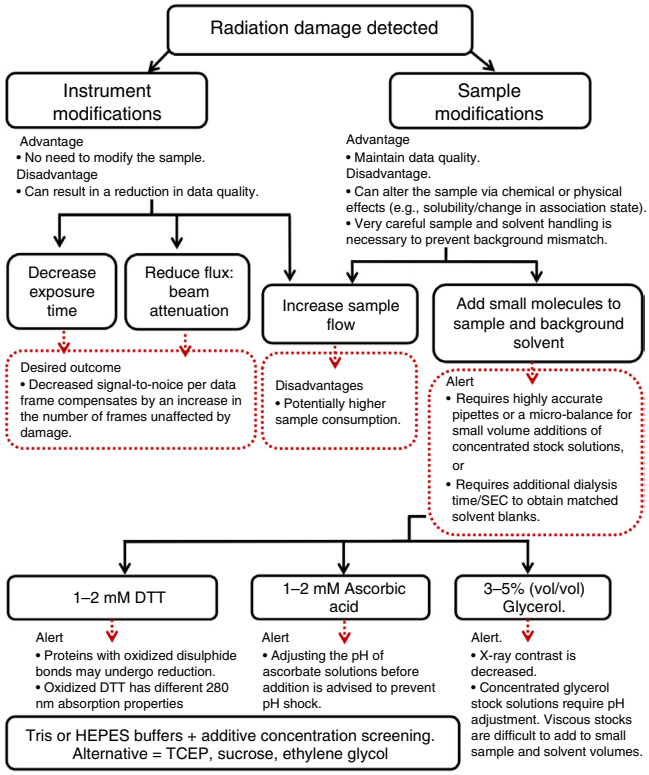


Figure 13 | Troubleshooting. Flowchart of options for reducing radiation damage to samples at a SAXS beam line. If X-ray radiation damage is detected in a sample, e.g., a systematic increase in $I(0)$ and R_g during the course of X-ray exposure is observed, instrument and/or sample modifications can be implemented to reduce its effects. Caution must be applied when adding small molecules (DTT, ascorbic acid and glycerol) to curb the effects of radiation damage. These small molecules must not radically alter a sample and cause structural changes or chemically induce aggregation. This can be tested before SAXS using DLS. It is also vital that equal quantities of small molecules are added to both the sample and the matched solvent blank in order to obtain the correct background subtraction required for the SAXS measurements.

? TROUBLESHOOTING

Limiting the effects of X-ray-induced aggregation

Several practical steps can be taken to reduce the effects of X-ray-induced aggregation of samples[56,57]. Radiation damage is particularly relevant at high-brilliance synchrotron SAXS beam lines, but it can also readily occur in samples exposed to lab-based X-ray sources. A general hypothesis is that X-ray radiation damage is caused by the photolysis of water into hydroxyl, hydroperoxyl radicals and solvated electrons. These highly reactive species can cause proteins to self-associate and irreversibly aggregate. Radiation damage can be detected in SAXS data by the following steps:

(i) Monitor the increase in $I(0)$ and R_g in the processed SAXS data.

(ii) Directly compare the full $I(q)$ versus q SAXS profiles of unsubtracted (or subtracted) data frames and identify significant differences in the intensities between frames at low-angles—e.g., using the Correlation Map method[93].

Refer to the troubleshooting flowchart in **Figure 13** for details on what steps can be taken to curb the effects of radiation damage.

● TIMING

Section 1, Steps 1–6, SAS sample purity, quality and preparation of the solvent blank: The production and characterization of purified macromolecules for SAS can span a few days to several weeks. SDS–PAGE analysis takes 1–2 h per gel, including setup, staining and destaining. Plan on longer times for native PAGE, e.g., up to 4 h per gel. SEC, SEC-MALLS or SEC-RALLS requires several hours to 1 d to complete, including buffer preparation, column equilibration, instrument calibration and performing the sample runs. Save time by equilibrating columns and detectors overnight. For SEC–SAXS, refer to **Box 1**. Stand-alone DLS and SLS require 30 min of instrument equilibration time, and each measurement takes ~100 s (e.g., 10 × 10 s each). Set aside 1 d to perform the requisite DLS tests on the sample. Sample dialysis typically requires 1 h of setup time and overnight buffer exchange. Centrifugation-based protein concentration may take 30–60 min to complete, depending on the selected speed and final desired concentration of the sample.

Section 2, Steps 7 and 8, quantity guides, sample concentration and MW: For polynucleotide and/or protein concentration determination, 1–2 min per sample is required (which includes blanking the spectrophotometer). For MW estimates using SAXS, set aside time to constitute the MW standard (if used) in the appropriate matched solvent—e.g., lysozyme requires overnight dialysis for buffer exchange. The time required for regular SAXS measurements, which includes loading the sample and solvent, collecting the sample and solvent data, and washing and drying the sample cell/capillary between measurements, takes anywhere between 2–10 min (high-flux synchrotron SAXS) and 1–4 h (lab-based X-ray source). Calculating the contrast using MULCh takes ~5 min (**Box 2**).

Section 3, Steps 9–35, preparation of samples for SANS experiments: calculating the neutron- scattering contrasts of a sample: It can take 4–6 weeks (or longer) to prepare for a SANS experiment, including protein purification and solubility/stability testing. ^{2}H-labeling may require additional time; refer to **Box 3**. Importantly, it is advised to plan the timing of a SANS at a nuclear facility before arrival. Day 1: devote to security and radiation safety training (1–5 h), as well as setting up samples for overnight dialysis (6–7 h) and cleaning the sample cells/quartz cuvettes. Day 2–*x*: perform the SANS measurements, typically 2–5 min transmissions for each sample and solvent + 10 min for blocked-beam measurements + 30 min for empty-cell measurements + 30–60 min for each solvent and sample (depending on counting statistics) + time for detector movements and the remeasurement of the empty cell, samples and buffers at any

additional detector positions. Day *x*: if the samples have been exposed to neutrons, leave time for the radiation safety team to check and clear the samples (15 min–1 h) before completing exit protocols and leaving the facility.
Box 1, in-line SEC–SAXS: buffer preparation + column and detector equilibration, 2–12 h; 1 × SEC–SAXS run: 30 min–2 h, depending on the SAXS beam line and SEC column flow rates; data processing: 20 min–2 h.
Box 2, calculation of X-ray and neutron-scattering contrasts using the *Contrast* module of MULCh: 5 min.
Box 3, deuteration of recombinant proteins using a laboratory-based ^{2}H labeling protocol: 5–10 days + additional time if screening of bacterial growth conditions is necessary.

ANTICIPATED RESULTS

Visit the Small-Angle Scattering Biological Data Bank (SASBDB; http://www.sasbdb.org) to view the results derived from solution SAS investigations. The SASBDB is a recent open-access initiative for the public dissemination of SAS data and modeling[17]. The fully searchable database was developed as part of a newly conceived federated database system (e.g., with Bioisis; http://www.bioisis.net) that incorporates recommendations from the wwPDB Small-angle Scattering Taskforce[63] (see also http://www.sasbdb.org/aboutSASBDB/ and http://www.sasbdb.org/help/). For example, refer to SASDBJ3 and SASDBK3 for the results obtained from SEC–SAXS experiments performed on BSA (**Box 1**; **Fig. 14**).

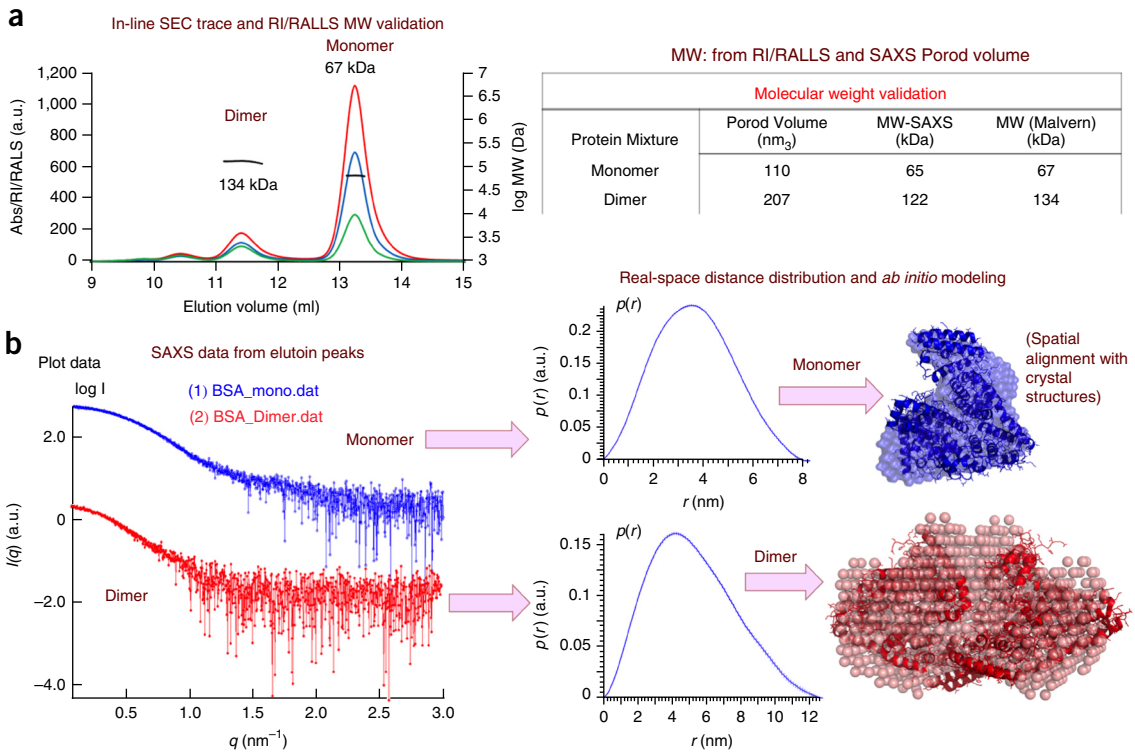


Figure 14 | SEC–SAXS component separation. (**a**) Results obtained from the separation of BSA monomers, dimers and higher-molecular weight (MW) oligomers using SEC with UV/RI/RALLS MW validation (Malvern Instruments 305 TDA detector). These data are measured in parallel with (**b**) the SAXS data collected as the BSA monomer and dimers elute from the SEC column. *Ab initio* bead models of the BSA monomer (blue surface) and dimer (red spheres) derived from the SAXS data are shown and compared with the published crystal structures of BSA (Protein Databank (PDB) 3V03). The SEC–SAXS/TDA data were collected at the EMBL-P12 BioSAXS beam line (DESY, Hamburg), and the results have been deposited into the Small-Angle Scattering Biological Data Bank, SASBDB (SASDBJ3 and SASDBK3).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS This work was supported by the Bundesministerium für Bildung und Forschung (BMBF) project BIOSCAT, grant 05K12YE1, by the European Community' Seventh Framework Programme (FP7/2007-2013) under BioStruct-X (grant agreement no. 283570) and by an HFSP grant (RGP0017/2012 to D.I.S. and C.M.J.). M.A.G. was supported by an EMBL Interdisciplinary Postdoc Programme (EIPOD) and Marie Curie COFUND actions. We thank J. Trewhella, in whose laboratory many of the procedures were performed. We also thank A. Duff (Australian Nuclear Science and Technology Organisation (ANSTO)) and David Jacques (LMB, Cambridge) for constructive comments on the deuteration incorporation spreadsheet.

AUTHOR CONTRIBUTIONS C.M.J., M.A.G., C.E.B., D.B.L., A.E.W. and D.I.S. helped develop SAXS and SANS sample preparation protocols and analytical tools. C.M.J., M.A.G., C.E.B. and D.I.S. performed radiation damage studies and developed protocols for SEC–SAXS. C.M.J., A.E.W. and D.B.L. contributed to 'in-house' ^{2}H-labeling protocols. D.B.L., A.E.W., C.M.J. and D.I.S. optimized protocols for preparing samples for SANS with contrast variation. A.E.W. developed *Contrast*. C.M.J., M.A.G., C.E.B., D.B.L., A.E.W. and D.I.S. critically discussed and wrote the manuscript.

COMPETING FINANCIAL INTERESTS The authors declare no competing financial interests. Although particular commercial products are noted throughout the text, such references should not be interpreted as product endorsements.

1. Graewert, M.A. & Svergun, D.I. Impact and progress in small and wide angle X-ray scattering (SAXS and WAXS). *Curr. Opin. Struct. Biol.* **23**, 748–754 (2013).
2. Hura, G.L. *et al.* Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nat. Meth.* **6**, 606–612 (2009).
3. Skou, S., Gillilan, R.E. & Ando, N. Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nat. Protoc.* **9**, 1727–1739 (2014).
4. Petoukhov, M.V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J. App. Crystallogr.* **45**, 342–350 (2012).
5. Petoukhov, M.V., Konarev, P.V., Kikhney, A.G. & Svergun, D.I. ATSAS 2.1 - towards automated and web-supported small-angle scattering data analysis. *J. Appl. Crystallogr.* **40**, s223–s228 (2007).
6. Franke, D., Kikhney, A.G. & Svergun, D.I. Automated acquisition and analysis of small angle X-ray scattering data. *Nuc. Instr. Meth. Phys. Res. Sec. A* **689**, 52–59 (2012).
7. Toft, K.N. *et al.* High-throughput small angle X-ray scattering from proteins in solution using a microfluidic front-end. *Anal. Chem.* **80**, 3648–3654 (2008).
8. Blanchet, C.E. *et al.* Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J. Appl. Crystallogr.* **48**, 431–443 (2015).
9. Petoukhov, M.V. & Svergun, D.I. Applications of small-angle X-ray scattering to biomacromolecular solutions. *Int. J. Biochem. Cell Biol.* **45**, 429–437 (2013).
10. Koch, M.H., Vachette, P. & Svergun, D.I. Small-angle scattering: a view on the properties, structures and structural changes of biological macromolecules in solution. *Q Rev. Biophys.* **36**, 147–227 (2003).
11. Wall, M.E., Gallagher, S.C. & Trewhella, J. Large-scale shape changes in proteins and macromolecular complexes. *Ann. Rev. Phys. Chem.* **51**, 355–380 (2000).
12. Mertens, H.D. & Svergun, D.I. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol.* **172**, 128–141 (2010).
13. Rambo, R.P. & Tainer, J.A. Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering. *Curr. Opin. Struct. Biol.* **20**, 128–137 (2010).
14. Krueger, S., Shin, J.H., Raghunandan, S., Curtis, J.E. & Kelman, Z. Atomistic ensemble modeling and small-angle neutron scattering of intrinsically disordered protein complexes: applied to minichromosome maintenance protein. *Biophys. J.* **101**, 2999–3007 (2011).
15. Whitten, A.E. *et al.* The structure of the KinA-Sda complex suggests an allosteric mechanism of histidine kinase inhibition. *J. Mol. Biol.* **368**, 407–420 (2007).
16. Vestergaard, B. *et al.* A helical structural nucleus is the primary elongating unit of insulin amyloid fibrils. *PLoS Biol.* **5**, e134 (2007).
17. Valentini, E., Kikhney, A.G., Previtali, G., Jeffries, C.M. & Svergun, D.I. SASBDB, a repository for biological small-angle scattering data. *Nucleic Acids Res.* **43**, D357–D363 (2015).
18. Franke, D. & Svergun, D.I. DAMMIF, a program for rapid *ab-initio* shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42**, 342–346 (2009).
19. Svergun, D.I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* **76**, 2879–2886 (1999).
20. Svergun, D.I., Petoukhov, M.V. & Koch, M.H. Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.* **80**, 2946–2953 (2001).
21. Petoukhov, M.V. & Svergun, D.I. Global rigid body modeling of macromolecular complexes against small-angle scattering data. *Biophys. J.* **89**, 1237–1250 (2005).
22. Sali, A. *et al.* Outcome of the First wwPDB Hybrid/Integrative Methods Task Force Workshop. *Structure* **23**, 1156–1167 (2015).
23. Panjkovich, A. & Svergun, D.I. Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis. *Phys. Chem. Chem. Phys.* **18**, 5707–5719 (2015).
24. Hammel, M. Validation of macromolecular flexibility in solution by small-angle X-ray scattering (SAXS). *Eur. Biophys. J.* **41**, 789–799 (2012).
25. Bernado, P., Mylonas, E., Petoukhov, M.V., Blackledge, M. & Svergun, D.I. Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem Soc.* **129**, 5656–5664 (2007).
26. Wang, Y., Trewhella, J. & Goldenberg, D.P. Small-angle X-ray scattering of reduced ribonuclease A: effects of solution conditions and comparisons with a computational model of unfolded proteins. *J. Mol. Biol.* **377**, 1576–1592 (2008).
27. Goldenberg, D.P. & Argyle, B. Minimal effects of macromolecular crowding on an intrinsically disordered protein: a small-angle neutron scattering study. *Biophys. J.* **106**, 905–914 (2014).
28. Tria, G., Mertens, H.D., Kachala, M. & Svergun, D.I. Advanced ensemble modelling of flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207–217 (2015).
29. Kikhney, A.G. & Svergun, D.I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* **589**, 2570–2577 (2015).
30. Kachala, M., Valentini, E. & Svergun, D.I. Application of SAXS for the Structural Characterization of IDPs. *Adv. Exp. Med. Biol.* **870**, 261–289 (2015).
31. Jimenez-Garcia, B., Pons, C., Svergun, D.I., Bernado, P. & Fernandez-Recio, J. pyDockSAXS: protein-protein complex structure by SAXS and computational docking. *Nucleic Acids Res.* **43**, W356–W361 (2015).
32. Cho, H.S., Schotte, F., Dashdorj, N., Kyndt, J. & Anfinrud, P.A. Probing anisotropic structure changes in proteins with picosecond time-resolved small-angle X-ray scattering. *J. Phys. Chem. B* **117**, 15825–15832 (2013).
33. Graceffa, R. *et al.* Sub-millisecond time-resolved SAXS using a continuous-flow mixer and X-ray microbeam. *J. Synchrotron Radiat.* **20**, 820–825 (2013).
34. Roessle, M. *et al.* Time-resolved small angle scattering: kinetics and structural data from proteins in solution. *J. Appl. Crystallogr.* **33**, 548–551 (2000).
35. Jacques, D.A. & Trewhella, J. Small-angle scattering for structural biology--expanding the frontier while avoiding the pitfalls. *Protein Sci.* **19**, 642–657 (2010).
36. Grishaev, A. Sample preparation, data collection and preliminary data analysis in biomolecular solution X-ray scattering. *Curr. Protoc. Protein Sci.* **70**, 17.14.1–17.14.18 (2012).
37. Glatter, O. & Kratky, O. *Small Angle X-ray Scattering* (Academic Press, 1982).
38. Feigin, L.A. & Svergun, D.I. *Structure Analysis by Small-angle X-ray and Neutron Scattering* (Plenum Press, 1987).
39. Svergun, D.I., Koch, M.H.J., Timmins, P.A. & May, R.P. *Small Angle X-Ray and Neutron Scattering From Solutions of Biological Macromolecules.* 1st edn, (Oxford University Press, 2013).
40. Jacques, D.A., Guss, J.M., Svergun, D.I. & Trewhella, J. Publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution. *Acta Crystallogr. Sec.* **D68**, 620–626 (2012).
41. Svergun, D.I. & Nierhaus, K.H. A map of protein-rRNA distribution in the 70 S *Escherichia coli* ribosome. *J. Biol. Chem.* **275**, 14432–14439 (2000).
42. Whitten, A.E., Jeffries, C.M., Harris, S.P. & Trewhella, J. Cardiac myosin-binding protein C decorates F-actin: implications for cardiac function. *Proc. Natl. Acad. Sci. USA* **105**, 18360–18365 (2008).
43. Lakey, J.H. Neutrons for biologists: a beginner's guide, or why you should consider using neutrons. *J. R. Soc.* **6**, S567–S573 (2009).
44. Zaccai, G. Straight lines of neutron scattering in biology: a review of basic controls in SANS and EINS. *Eur. Biophys. J.* **41**, 781–787 (2012).
45. Krueger, S. SANS provides unique information on the structure and function of biological macromolecules in solution. *Physica B* **241-243**, 1131–1137 (1997).
46. Heller, W.T. & Littrell, K.C. Small-angle neutron scattering for molecular biology: basics and instrumentation. *Meth. Mol. Biol.* **544**, 293–305 (2009).
47. Krueger, J.K. & Wignall, G.D. Small-Angle Neutron Scattering from Biological Molecules. in *Neutron Scattering in Biology* 127–160 (Berlin/Heidelberg: Springer, 2006).
48. Svergun, D.I. & Koch, M.H.J. Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.* **66**, 1735–1782 (2003).
49. Jacrot, B. & Zaccai, G. Determination of molecular weight by neutron scattering. *Biopolymers* **20**, 2413–2426 (1981).
50. Sears, V.F. Neutron scattering lengths and cross sections. *Neutron News* **3**, 26–37 (1992).
51. Whitten, A.E., Cai, S. & Trewhella, J. MULCh: modules for the analysis of small-angle neutron contrast variation data from biomolecular assemblies. *J. App. Crystallogr.* **41**, 222–226 (2008).
52. Konarev, P.V. & Svergun, D.I. A posteriori determination of the useful data range for small-angle scattering experiments on dilute monodisperse systems. *IUCrJ* **2**, 352–360 (2015).
53. Lafleur, J.P. *et al.* Automated microfluidic sample-preparation platform for high-throughput structural investigation of proteins by small-angle X-ray scattering. *J. App. Crystallogr.* **44**, 1090–1099 (2011).

54. Skou, M., Skou, S., Jensen, T.G., Vestergaard, B. & Gillilan, R.E. icrofluidic dialysis for biological small-angle X-ray scattering. *J. Appl. Crystallogr.* **47**, 1355–1366 (2014).
55. Bertani, G. Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* **62**, 293–300 (1951).
56. Kuwamoto, S., Akiyama, S. & Fujisawa, T. Radiation damage to a protein solution, detected by synchrotron X-ray small-angle scattering: dose-related considerations and suppression by cryoprotectants. *J. Synchrotron Radiat.* **11**, 462–468 (2004).
57. Jeffries, C.M., Graewert, M.A., Svergun, D.I. & Blanchet, C.E. Limiting radiation damage for high brilliance biological solution scattering: practical experience at the EMBL P12 beamline, PETRAIII. *J. Synchrotron Radiat.* **22**, 273–279 (2014).
58. Folta-Stogniew, E. & Williams, K.R. Determination of molecular masses of proteins in solution: implementation of an HPLC size exclusion chromatography and laser light scattering service in a core laboratory. *J. Biomol. Tech.* **10**, 51–63 (1999).
59. Jacques, D.A. *et al.* Structure of the sporulation histidine kinase inhibitor Sda from *Bacillus subtilis* and insights into its solution state. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 574–581 (2009).
60. Dünweg, B., Reith, D., Steinhauser, M. & Kremer, K. Corrections to scaling in the hydrodynamic properties of dilute polymer solutions. *J. Chem. Phys.* **117**, 914–924 (2002).
61. Brewer, A.K. & Striegel, A.M. Characterizing the size, shape, and compactness of a polydisperse prolate ellipsoidal particle via quadruple-detector hydrodynamic chromatography. *Analyst* **136**, 515–519 (2011).
62. Rubinson, K.A., Stanley, C. & Krueger, S. Small-angle neutron scattering and the errors in protein structures that arise from uncorrected background and intermolecular interactions. *J. App. Crystallogr.* **41**, 456–465 (2008).
63. Trewhella, J. *et al.* Report of the wwPDB Small-Angle Scattering Task Force: data requirements for biomolecular modeling and the PDB. *Structure* **21**, 875–881 (2013).
64. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
65. Gasteiger, E. *et al.* Protein identification and analysis tools on the ExPASy server. in *The Proteomics Protocols Handbook* (ed. Walker, J.M.) 571–607 (Humana Press, 2005).
66. Cleland,, W.W. Dithiothreitol, a new protective reagent for sh groups. *Biochemistry* **3**, 480–482 (1964).
67. Stevens, R., Stevens, L. & Price, N.C. The stabilities of various thiol compounds used in protein purifications. *Biochem. Educ.* **11**, 70–70 (1983).
68. Krežel, A., Latajka, R., Bujacz, G.D. & Bal, W. Coordination properties of Tris(2-carboxyethyl)phosphine, a newly introduced thiol reductant, and its oxide. *Inorg. Chem.* **42**, 1994–2003 (2003).
69. Krezel, A. *et al.* Coordination of heavy metals by dithiothreitol, a commonly used thiol group protectant. *J. Inorg. Biochem.* **84**, 77–88 (2001).
70. Han, J.C. & Han, G.Y. A procedure for quantitative determination of Tris(2-Carboxyethyl)phosphine, an odorless reducing agent more stable and effective than dithiothreitol. *Anal. Biochem.* **220**, 5–10 (1994).
71. Zhao, H., Brown, P.H. & Schuck, P. On the distribution of protein refractive index increments. *Biophys. J.* **100**, 2309–2317 (2011).
72. Guinier, A. La diffraction des rayons X aux très pétits angles; application á l'etude de phénomènes ultramicroscopiques. *Ann. Phys.* **12**, 161–237 (1939).
73. Svergun, D.I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495–503 (1992).
74. Glatter, O. A new method for the evaluation of small-angle scattering data. *J. Appl. Crystallogr.* **10**, 415–421 (1977).
75. Mylonas, E. & Svergun, D.I. Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *J. Appl. Crystallogr.* **40**, s245–s249 (2007).
76. Krigbaum, W.R. & Kugler, F.R. Molecular conformation of egg-white lysozyme and bovine alpha-lactalbumin in solution. *Biochemistry* **9**, 1216–1223 (1970).
77. Watson, M.C. & Curtis, J.E. Probing the average local structure of biomolecules using small-angle scattering and scaling laws. *Biophys. J.* **106**, 2474–2482 (2014).
78. Fischer, H., de Oliveira Neto, M., Napolitano, H.B., Polikarpov, I. & Craievich, A.F. Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale. *J. Appl. Crystallogr.* **43**, 101–109 (2010).
79. Rambo, R.P. & Tainer, J.A. Accurate assessment of mass, models and resolution by small-angle scattering. *Nature* **496**, 477–481 (2013).
80. Fritz, G. & Bergmann, A. SAXS instruments with slit collimation: investigation of resolution and flux. *J. Appl. Crystallogr.* **39**, 64–71 (2006).
81. Svergun, D.I., Barberato, C. & Koch, M.H.J. CRYSOL - a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773 (1995).
82. Voss, N.R. & Gerstein, M. Calculation of standard atomic volumes for RNA and comparison with proteins: RNA is packed more tightly. *J. Mol. Biol.* **346**, 477–492 (2005).
83. Orthaber, D., Bergmann, A. & Glatter, O. SAXS experiments on absolute scale with Kratky systems using water as a secondary standard. *J. Appl. Crystallogr.* **33**, 218–225 (2000).
84. Wignall, G.D. & Bates, F.S. Absolute calibration of small-angle neutron scattering data. *J. App. Crystallogr.* **20**, 28–40 (1987).
85. Svergun, D.I. *et al.* Protein hydration in solution: experimental observation by x-ray and neutron scattering. *Proc. Natl. Acad. Sci. USA* **95**, 2267–2272 (1998).
86. Leiting, B., Marsilio, F. & O'Connell, J.F. Predictable deuteration of recombinant proteins expressed in *Escherichia coli*. *Anal. Biochem.* **265**, 351–355 (1998).
87. Goryunov, A.S. H/D isotope effects on protein hydration and interaction in solution. *Gen. Physiol. Biophys.* **25**, 303–311 (2006).
88. Sheu, S.Y., Schlag, E.W., Selzle, H.L. & Yang, D.Y. Molecular dynamics of hydrogen bonds in protein-D2O: the solvent isotope effect. *J. Phys. Chem. A* **112**, 797–802 (2008).
89. Baghurst, P.A., Nichol, L.W. & Sawyer, W.H. The effect of D 2 O on the association of -lactoglobulin A. *J. Biol. Chem.* **247**, 3199–3204 (1972).
90. Dougan, L., Koti, A.S., Genchev, G., Lu, H. & Fernandez, J.M. A single-molecule perspective on the role of solvent hydrogen bonds in protein folding and chemical reactions. *Chemphyschem* **9**, 2836–2847 (2008).
91. Tehei, M., Madern, D., Pfister, C. & Zaccai, G. Fast dynamics of halophilic malate dehydrogenase and BSA measured by neutron scattering under various solvent conditions influencing protein stability. *Proc. Natl. Acad. Sci. USA* **98**, 14356–14361 (2001).
92. Jasnin, M., Tehei, M., Moulin, M., Haertlein, M. & Zaccai, G. Solvent isotope effect on macromolecular dynamics in *E. coli*. *Eur. Biophys. J.* **37**, 613–617 (2008).
93. Graewert, M.A. *et al.* Automated pipeline for purification, biophysical and x-ray analysis of biomacromolecular solutions. *Sci. Rep.* **5**, 10734 (2015).
94. Glasoe, P.K. & Long, F.A. Use of glass electrodes to measure acidities in deuterium oxide. *J. Phys. Chem.* **64**, 188–190 (1960).
95. Svergun, D.I., Koch, M.H.J., Timmins, P.A. & May, R.P. *Small Angle X-Ray and Neutron Scattering from Solutions of Biological Macromolecules*. (Oxford University Press, 2013).
96. Wood, K. *et al.* Exploring the structure of biological macromolecules in solution using Quokka, the small angle neutron scattering instrument, at ANSTO. *Nucl. Instrum. Meth. A* **798**, 44–51 (2015).
97. Mokbel, N. *et al.* K7del is a common TPM2 gene mutation associated with nemaline myopathy and raised myofibre calcium sensitivity. *Brain* **136**, 494–507 (2013).
98. Mathew, E., Mirza, A. & Menhart, N. Liquid-chromatography-coupled SAXS for accurate sizing of aggregating proteins. *J. Synchrotron. Radiat.* **11**, 314–318 (2004).
99. David, G. & Perez, J. Combined sampler robot and high-performance liquid chromatography: a fully automated system for biological small-angle X-ray scattering experiments at the Synchrotron SOLEIL SWING beamline. *J. App. Crystallogr.* **42**, 892–900 (2009).
100. Chen, X. *et al.* High yield expression and efficient purification of deuterated human protein galectin-2. *Food Bioprod. Process.* **90**, 563–572 (2012).
101. Sun, G. *et al.* Expression, purification and crystallization of human kynurenine aminotransferase 2 exploiting a highly optimized codon set. *Protein Expr. Purif.* **121**, 41–45 (2016).