



2023 update of template tables for reporting biomolecular structural modelling of small-angle scattering data

Jill Trehella,^{a*} Cy M. Jeffries^b and Andrew E. Whitten^c

Received 14 November 2022

Accepted 23 December 2022

Edited by C. S. Bond, University of Western Australia, Crawley, Australia

Keywords: small-angle scattering; SAXS; SANS; contrast variation; biomolecular structural modelling; template tables.

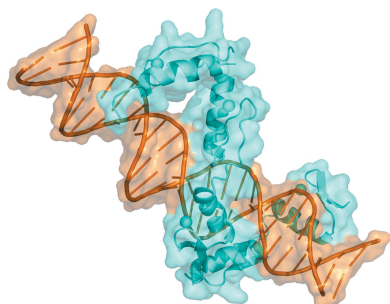
Supporting information: this article has supporting information at journals.iucr.org/d

^aSchool of Life and Environmental Sciences, The University of Sydney, Sydney, NSW 2006, Australia, ^bEuropean Molecular Biology Laboratory (EMBL), Hamburg Unit, Notkestrasse 85, c/o Deutsches Elektronen-Synchrotron, 22607 Hamburg, Germany, and ^cAustralian Nuclear Science and Technology Organisation, New Illawarra Road, Lucas Heights, NSW 2234, Australia. *Correspondence e-mail: jill.trehella@sydney.edu.au

In 2017, guidelines were published for reporting structural modelling of small-angle scattering (SAS) data from biomolecules in solution that exemplified best-practice documentation of experiments and analysis. Since then, there has been significant progress in SAS data and model archiving, and the IUCr journal editors announced that the IUCr biology journals will require the deposition of SAS data used in biomolecular structure solution into a public archive, as well as adherence to the 2017 reporting guidelines. In this context, the reporting template tables accompanying the 2017 publication guidelines have been reviewed with a focus on making them both easier to use and more general. With input from the SAS community via the IUCr Commission on SAS and attendees of the triennial 2022 SAS meeting (SAS2022, Campinas, Brazil), an updated reporting template table has been developed that includes standard descriptions for proteins, glycosylated proteins, DNA and RNA, with some reorganization of the data to improve readability and interpretation. In addition, a specialized template has been developed for reporting SAS contrast-variation (SAS-cv) data and models that incorporates the additional reporting requirements from the 2017 guidelines for these more complicated experiments. To demonstrate their utility, examples of reporting with these new templates are provided for a SAS study of a DNA–protein complex and a SAS-cv experiment on a protein complex. The examples demonstrate how the tabulated information promotes transparent reporting that, in combination with the recommended figures and additional information best presented in the main text, enables the reader of the work to readily draw their own conclusions regarding the quality of the data and the validity of the models presented.

1. Introduction

The 2017 publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution (Trehella *et al.*, 2017) aimed to provide a reporting framework for publication that would enable reviewers and readers to independently assess the quality of the data and the interpretations made by the authors. A major motivation of the effort to establish this community consensus for what should be presented was the desire to ensure that biomolecular SAS achieved its full potential, especially in the emerging field of integrative/hybrid structure determination (Grishaev, 2017; Sali *et al.*, 2015; Brosey & Tainer, 2019; Sali, 2021; Schneidman-Duhovny *et al.*, 2012; Schroer & Svergun, 2018; Trehella, 2016, 2022). When SAS data are used in conjunction with other methods to solve the structures of complex assemblies of biological importance, especially where atomistic detail is presented, it is critical that there are standards for data-quality assurance along with agreed tools and



OPEN ACCESS

Published under a CC BY 4.0 licence

methods for model validation. This ensures that the broader structural biology community has confidence in a result that becomes part of the public archive of structures and is subsequently integrated across the numerous data and bioinformatic resources for structural biology (Vallat *et al.*, 2018, 2021; Berman *et al.*, 2020; Berman, Adams *et al.*, 2018; Berman, Trewhella *et al.*, 2018; Burley *et al.*, 2017).

The recommendations in the 2017 guidelines were themselves an update that built upon earlier work undertaken with community engagement via the International Union of Crystallography (IUCr) Commissions for SAS (CSAS) and Journals, discussions at the triennial SAS meetings and the worldwide Protein Data Bank (wwPDB) through its SAS validation task force (SASvtf) (Trewhella *et al.*, 2013; Jacques, Guss & Trewhella, 2012; Jacques, Guss, Svergun *et al.*, 2012). The guidelines are comprehensive regarding quality assurance in sample preparation and characterization, data acquisition and processing, analysis and modelling. Further, they include the recommendation for data and associated models to be deposited in a public archive. Today, the Small Angle Scattering Biological Data Bank (SASBDB; Valentini *et al.*, 2015) contains >3000 experimental data sets and >4000 associated models. With these developments (reviewed in Berman, Adams *et al.*, 2018; Trewhella, 2018), the wwPDB deposition system (OneDep; Young *et al.*, 2017) initiated streamlining via an API (application programming interface) with the SASBDB so that, for the first time, SAXS data supporting NMR/SAXS structures would be available to reviewers and published to the broader community (Kikhney *et al.*, 2020). Since then, SAS data have been formally included in the prototype archive for integrative/hybrid methods PDB-Dev (Vallat *et al.*, 2021), where the overall model validation report under development includes assessments of SAS data quality and model fit. The SAS data are linked to their depositions in the SASBDB, where sliding-scale indicator metrics for data validation and data–model fitting are used to give a visual summary of quality assessments.

As proponents of the FAIR (Findable, Accessible, Interoperable and Reusable; Wilkinson *et al.*, 2016) and FACT (Fair, Accurate, Confidential and Transparent; van der Aalst *et al.*, 2017; Helliwell, 2019) principles of publishing and in the context of the developments described above, the IUCr journals announced that for their biology journals (*Acta Crystallographica Section D*, *Acta Crystallographica Section F* and *IUCrJ*)

for any manuscript containing conventional structures determined by the most common techniques (crystallography, NMR, cryoEM, SAXS) the data that are deposited with the relevant database to obtain the accession code and validation reports must be uploaded prior to editorial review

(Baker *et al.*, 2022). Accompanying this requirement for data deposition, these journals ‘are unifying their requirements for information to be included in the standard tables for various experiment types (colloquially Table 1 for crystal structures). These new requirements will be described in detail in the revised Notes for Authors for the journals’. For SAS data,

these notes specify that ‘For structures determined by small-angle scattering, authors should deposit their data at the SASBDB (<https://www.sasbdb.org/>) in advance of submission to the journal and provide the SASBDB reference code(s)’. In addition, ‘Authors should follow the publication guidelines’ from Trewhella *et al.* (2017).

With five years of experience using the template tables developed as part of the 2107 publication guidelines and the substantial growth in SASBDB, it was timely to review the template tables to ensure that they were achieving the original goal without imposing unnecessary work on researchers. Practical experience indicated that some reorganization of the information would benefit authors and readers alike. Further, the original template table drew largely on example small-angle X-ray scattering (SAXS) experiments from four proteins as presented in the paper. The template tables thus have been reviewed and updated considering the guidelines in the context of more complex samples and for SAS contrast-variation (SAS-cv) experiments that most often include small-angle neutron scattering (SANS) data. Importantly, with this update we do not purport to change the content of what constitutes ‘best practice’ in documenting biomolecular SAS data that are used for 3D structure modelling as described in the 2017 guidelines, but rather simply to improve the presentation. Two illustrative examples are provided to demonstrate use of the templates, which are analysed largely, although not exclusively, using the *ATSAS* software, which is popular with biomolecular SAS users due to a combination of ease of access and broad utility. However, there are numerous alternative packages that offer users different features and use different methods. As an aid to new SAS users, we therefore have provided as comprehensive a list as we could assemble at this time in the supporting information with indicators of their utility (for example imaging, data reduction, analysis, modelling) and references (Supplementary Table S1) along with useful related links (Supplementary Table S2).

2. Process for updating the template table

Input on potential revisions to the guidelines was first requested via email from members of the IUCr CSAS, the SASvtf and the 46 co-authors of the recently published benchmarking study for biomolecular modelling (Trewhella *et al.*, 2022). Subsequent in-person discussions at the open meeting of the IUCr CSAS at the 18th International Small-Angle Scattering Conference (SAS2022; Campinas, Brazil, 11–16 September 2022) led to the formation of a small working group to (i) consider, in light of the input received, whether information could be better organized within the template table to achieve greater clarity and whether some information would be better described elsewhere, for example in the experimental methods text, (ii) generalize the sample descriptions, explicitly including nonprotein components, for example DNA/RNA and glycans/carbohydrates, (iii) develop a SAS-cv template that includes the additional recommendations from the 2017 guidelines for this class of experiment and (iv) test the revised templates with example data sets.

From this process, two template tables have been developed: (i) an updated template for the general biomolecular structural modelling SAS experiment and (ii) a specific SAS-cv template (Supplementary Tables S3 and S4, respectively). In reporting an experiment, the information captured using the appropriate template table would be accompanied by the additional information detailed in the 2017 guidelines, noting that the guidelines for figures to be presented include the following.

(i) The scattering profile $I(q)$ versus q (log-linear, plus log-log if useful).

(ii) A Kratky plot, with a preference for dimensionless Kratky plots [$(qR_g)^2 I(q)/I(0)$ versus qR_g] in most cases (with the exception being for SANS contrast-variation data, where negative R_g values are possible).

(iii) The pairwise distance distribution function [PDDF or $P(r)$ versus r].

(iv) For SEC-SAS data, R_g and $I(0)$ as a function of time or measurement frame number.

(v) Model fits shown as log-linear plots with the experimental data accompanied by residual error-weighted difference plots that complement the tabulated numerical parameters χ^2 and the Correlation Map (CorMap) P -value (Franke *et al.*, 2015).

3. Changes to the general template and their rationale

The main changes made to the original panels (a)–(f) for the general template (Supplementary Table S3) are as follows.

(i) Panel (a) *Sample details* now consolidates all sample information with recommended standard naming conventions.

(ii) Panel (b) *SAS data collection* is simplified by moving some information that is more suited to methods.

(iii) The original panel (c) that listed the software used is eliminated and this information is now co-located in the respective panels where their results are reported.

(iv) The new panel (c) *SAS-derived structural parameters* is essentially the same as the original panel (d) except the molecular mass from $I(0)$ and Porod volume, V_p , results are moved to the new panel (d).

(v) The new panel (d) *Scattering particle size* consolidates all information relating to size and standardizes to reporting molecular mass M in Daltons (or kDa).

(vi) Panel (e) *Modelling* is now a single panel where all structural modelling is consolidated.

(vii) Panel (f) *Data and model deposition* specifies deposition to SASBDB.

3.1. Standard nomenclature and public archive IDs for describing components

The template table recommends standard nomenclature for chemical groups and that where possible descriptions of the sequences of proteins, DNA, RNA and glycan components should reference their identifiers in public archives.

(i) For proteins, use the UniProt ID (<https://www.uniprot.org/>) and the recommended UniProt name with the amino-acid

sequence range of the construct measured by SAS, plus any tags, post-translational modifications, ligands, cofactors, metals *etc.* If UniProt IDs are not available the recommendation is to quote the NCBI accession and the protein name (<https://www.ncbi.nlm.nih.gov/guide/proteins/>). If a sequence has neither UniProt nor NCBI identifiers, or if the description is too long for the table format, provide an abbreviated title with a reference to the location where the exact sequences with modifications *etc.* can be found.

(ii) For DNA/RNA, if possible quote the relevant GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>), RNACentral (<https://rnacentral.org/>) or ENA (<https://www.ebi.ac.uk/ena/browser/home>) accession number, specifying any modifications, derivatives *etc.* If the description is too long for the table format, provide an abbreviated title with a reference to the location where the exact sequences with modifications *etc.* can be found.

(iii) For glycans, if possible quote the GlyTouCan (<https://glytoucan.org/>) accession code or information from GlyGen (<https://www.glygen.org/>). It is recommended to adhere to the Symbol Nomenclature for Glycans (SNFG) protocols (<https://pubmed.ncbi.nlm.nih.gov/31184695/>) and/or IUPAC nomenclature (<https://iupac.org/what-we-do/nomenclature/>).

3.2. Reorganized sample details and scattering particle size for sample-quality assurance

Because there are several methods for determining the molecular mass or volume of the scattering particle from experiment, there has been a tendency to report just one result. Panel (d) asks for the Porod volume V_p (calculated from the scattering invariant) and a concentration-dependent and a concentration-independent method for determining M from the SAS data, as well as any SAS-independent method. Consistency among these different methods and agreement with the calculated value based on chemical composition, within experimental error, is the strongest assurance that the experimental SAS profile is from the target of interest, unaffected by aggregation or interparticle correlations. Authors can address any anomalies among the experimentally determined M and V_p values, explaining whether there were problems with any one measurement or if assumptions do not hold. For example, the assumptions underlying the calculation of V_p do not hold for objects with non-uniform scattering densities or for partially unfolded objects. A reader or reviewer is then able to quickly evaluate their significance.

3.3. Changes in reporting data-collection parameters

The parameters for wavelength and beam geometry are no longer explicitly in the template table. It is not common practice to make these parameters routinely available to users at synchrotron SAXS facilities. While the information is normally in the headers of the raw data files, it is not always easy to find for the average user and in most cases SAXS data analysis can assume effective point geometry and a single-wavelength source. However, in the case of SAXS instruments using slit geometry or SANS measurements, the assumption of

effective point geometry does not hold. For SAXS instruments using slit geometry, the measured beam profile should be reported, while for SANS instruments the wavelength distribution, collimation lengths, source and sample aperture dimensions, detector distance and pixel sizes are important. These should either be reported in the methods or can be included in panel (b) *Data collection* as ‘Additional relevant details’, as demonstrated for the examples provided below.

3.4. Changes to reporting modelling parameters

Some of the information asked for in the original template relating to modelling protocols has not proven to easily fit into the table format. With the updated table, all details required to reproduce any modelling should be reported in the main text of the paper, including but not limited to adjustable fitting parameters. In the case of atomistic models, a full description of domain/subunit coordinates, their source, any regions of presumed flexibility and contacts used as constraints in rigid-body modelling should be reported in the main text.

3.5. SASBDB deposition

With its integration with the PDB, the major structural biology repository in the world, and now PDB-Dev, there is a clear advantage to depositing data with SASBDB. In addition, SASBDB provides quality assessment with metrics for data validation and data–model fitting. However, SASBDB is configured to accept and display one SAS profile per entry, which makes the deposition of SAS-cv data quite tedious. For this reason, SAS-cv data-series deposition is best achieved by requesting the inclusion of an ‘additional files’ folder (containing the complete SAS-cv series data; see SASBDB entry SASDZH3 described below for an example) along with

a representative profile from the series that would be displayed.

3.6. Testing the utility of the updated general template: a DNA–protein complex

To test for relative ease of use and for achieving the goal of transparency that would aid readers and reviewers alike, a published SAXS study reporting the structure of a protein–DNA complex was used to populate the updated general SAS template. The complex was made up of a zinc-finger protein (ZBTB38, a DNA-binding transcription regulator) and a double-methylated duplex DNA (mCZ38BS; Pozner *et al.*, 2018; Fig. 1*a*). SAXS data and associated models have been deposited in SASBDB (entries SASDCA3 and SASDCB3), and we note here that the model deposited as SASDCA3 has been updated since the publication of the 2018 paper based on the solution of a crystal structure (PDB entry 6e93) that contains major portions of the complex. Table 1 is obtained by populating the general SAS template with the SAXS data and the updated model, and it provides the reader or reviewer with a concise, comprehensive view of the samples measured and data-collection details, plus the results of the data analysis and modelling. Just some of the key things that are readily brought into focus are the following.

(i) Two samples were measured by SAXS; the doubly methylated 27-mer duplex DNA and its complex with the five-zinc-finger protein with five bound zinc ions.

(ii) The data were acquired on a laboratory-based instrument with a sealed-tube source and slit geometry, and details of how the slit smearing were handled are detailed in ‘Additional relevant details’ in panel (b) *SAS data collection*.

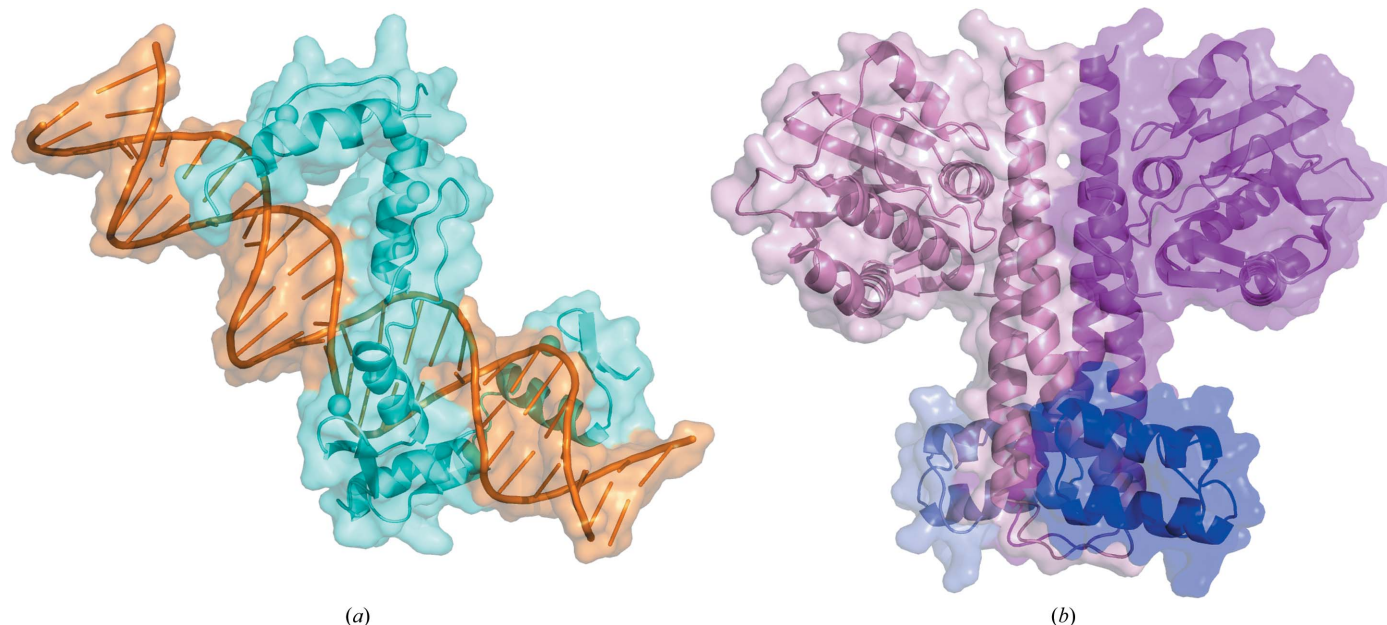


Figure 1

Depiction of the complexes for the SAS experiments used to illustrate the use of the template tables shown as cartoons with a transparent surface representation. (a) The zinc finger- and BTB domain-containing protein 38 (ZBTB38; cyan) complexed with methylated duplex DNA mCZ38BS (orange). (b) The histidine kinase KinA dimer (magenta and pink monomers) complexed with deuterated Sda (blue) (KinA₂–2^DSda).

Table 1

SAXS sample, data-collection, analysis and 3D modelling details for zinc finger- and BTB domain-containing protein 38 (ZBTB38) bound to the 27-mer doubly methylated DNA duplex (mCZ38BS).

(a) Sample details. DNA nucleotide sequences for mCZ38BS duplex DNA (with 5-methylcytosine, 5mC) are 5'-GCACTCAT(5mC)GG(5mC)GCAGATCAGC TAGCC-3' and 5'-GGCTAGCTGATCTG(5 mC)GC(5mC)GATGAGTGC-3'.

Organism	Human	
Source	Methylated DNA mCZ38BS, synthetic oligomers; ZBTB38 protein, <i>E. coli</i> (BL21) recombinant expression (Pozner <i>et al.</i> , 2018)	ZBTB38:mCZ38BS
Scattering particle composition	mCZ38BS	ZBTB38:mCZ38BS
Protein		Five-zinc-finger protein [ZBTB38, UniProt Q8NAP3(1006–1153)] with five bound zinc ions
DNA/RNA	27-mer duplex DNA with site-specific methylated cytosines (mCZ38BS), NCBI:txid9606	mCZ38BS
Stoichiometry of components	n.a.	1:1
Sample environment/configuration		
Solvent composition	10 mM Tris pH 6.8, 1 mM TCEP, 0.005% (w/v) NaN ₃ , 10% (v/v) D ₂ O	10 mM Tris pH 6.8, 1 mM TCEP, 0.005% (w/v) NaN ₃ , 10% (v/v) D ₂ O
Sample temperature (°C)	22	22
In-beam sample cell	1 mm quartz capillary, no flow	1 mm quartz capillary, no flow
Batch measurements		
Sample concentration(s) <i>c</i> (mg ml ⁻¹)	4.2	3.6

(b) SAS data collection.

Data-acquisition/reduction software	<i>SAXSquant</i> (Anton Paar)
Source/instrument description	Sealed-tube X-ray, Anton Paar SAXSess with Mythen 1D detector
Measured <i>q</i> -range (<i>q</i> _{min} – <i>q</i> _{max}) (Å ⁻¹)	0.008–0.7
Method for scaling intensities	Absolute scaling (cm ⁻¹) referenced to water
Exposure time(s), No. of exposures	30 s × 120 frames
Additional relevant details	10 mm line source with 262 mm sample-to-detector distance. Guinier analysis used <i>SAXSquant</i> desmeared data. The <i>GNOM P(r)</i> model was smeared (trapezoidal approximation for the slit geometry; AH/LH = 0.29/0.15 Å ⁻¹) for comparison with experiment. For modelling, <i>I(q)</i> was desmeared via point-by-point multiplication of the ratio of <i>I(q)</i> from the smeared <i>P(r)</i> to that from the unsmeared <i>P(r)</i> .

(c) SAS-derived structural parameters.

Method(s)/software	<i>PRIMUS</i> / <i>qt</i> , <i>AUTORG</i> and <i>GNOM</i> (<i>ATSAS</i> 2.6.0; Petoukhov <i>et al.</i> , 2012). mCZ38BS	ZBTB38:mCZ38BS
Guinier analysis		
<i>I</i> (0) ± σ (cm ⁻¹)	0.12 ± 0.002†	0.19 ± 0.003†
<i>R</i> _g ± σ (Å)	25.4 ± 0.7	24.3 ± 0.7
<i>qR</i> _g range (datapoint range)	0.25–1.27 (1–31)	0.24–1.28 (1–33)
Linear fit assessment (<i>AUTORG</i> fidelity)	0.99	0.80
PDDF/ <i>P(r)</i> analysis		
<i>I</i> (0) ± σ (cm ⁻¹)	0.114 ± 0.001	0.182 ± 0.002
<i>R</i> _g ± σ (Å)	26.15 ± 0.06	26.07 ± 0.03
<i>d</i> _{max} (Å)	92	82
<i>q</i> -range (Å ⁻¹)	0.010–0.372	0.081–0.362
<i>P(r)</i> reciprocal-space fit: χ ² , CorMap <i>P</i> -value	1.12, 0.85	1.44, 0.86

(iii) There is reasonable agreement between Guinier- and *P(r)*-derived parameters for both samples (for details regarding this criterion, see Section 2.3 of Trewhella *et al.*, 2017).

(iv) The fidelity of the linear fit to the Guinier region is excellent as assessed by *AUTORG* (Petoukhov *et al.*, 2012).

(v) The slit-smeared *P(r)*-model fit to the data is excellent when evaluated by the CorMap *P*-values, while the χ² values are >1 (1.12 and 1.44 for the DNA and complex, respectively.)

(vi) The molecular-weight determination from *I*(0)/*c* for the DNA and the complex are 9% lower and 25% higher than the expected values based on chemical composition, respectively.

(vii) The ratios of *V*_p to calculated *M* are in an acceptable range for a hydrated biomolecule in solution (Trewhella *et al.*, 2017).

(viii) Dummy-atom and atomistic modelling results give excellent fits to the data as measured by CorMap *P*-values, but

again there are χ² values >1 that are highest for the atomistic model of the complex (1.65).

With this quick assessment from the table, the reader would want to interrogate the paper, and potentially the deposited data and models, to find answers to the following questions.

(i) Why was scattering from the DNA but not the protein used in *MONSA* modelling of the DNA–protein complex? *Answer:* The paper reports that the zinc-finger protein alone aggregated in solution. Further, its five domains connected by linkers would be expected to adjust their relative dispositions upon binding DNA. In contrast, the double-helical DNA showed neither aggregation nor the effects of interparticle correlations and its structure is not likely to be radically altered in the free and bound forms.

(ii) Are the molecular-weight values from *I*(0)/*c* reasonable? *Answer:* Yes, the concentration determinations used UV extinction coefficients at 260 and 280 nm, for which a 10%

Table 1 (continued)

(d) Scattering particle size.

Method(s)/software	<i>PRIMUS/qt</i> (ATSAS 3.1; Manalastas-Cantos <i>et al.</i> , 2021) for M from Bayesian inference and V_p , <i>MULCh</i> (Whitten <i>et al.</i> , 2008) for v , $\Delta\rho$ and M from chemical composition, equation (1) in Trehwella <i>et al.</i> (2017) for M from $I(0)/c$	
	mCZ38BS	ZBTB38:mCZ38BS
Volume estimates (\AA^3)		
Porod volume V_p (ratio to M)	21100 (1.27)	53000 (1.54)
Molecular mass M estimates (Da)		
From chemical composition	16632	34389
From c -independent method (Bayesian inference, range with % confidence)	15150–18350, 98%	35750–38950, 95%
From $I(0)/c$ (ratio to expected)	18334 (0.91)	43100 (1.25)
Partial specific volume v ($\text{cm}^3 \text{g}^{-1}$)	0.591	0.657 (ZBTB38 alone 0.718)
Contrast $\Delta\rho$ (10^{10}cm^{-2})	5.425	4.158 (ZBTB38 alone 3.182)
From SAS-independent measure	n.a.	n.a.

(e) Modelling.

Methods/software	Dummy-atom (<i>DAMMIN</i>), multiphase dummy-atom (<i>MONSA1.45</i>) and rigid-body modelling (<i>SASREF7</i>) (ATSAS 2.6.0; Petoukhov <i>et al.</i> , 2012)	
	mCZ38BS	ZBTB38:mCZ38BS
Shape modelling/software	<i>DAMMIN</i>	<i>MONSA1.45</i>
q -range for fit (\AA^{-1})	0.01–0.2	0.0126–0.2
Symmetry/anisotropy assumptions	$P1$	$P1$
No. of individual model reconstructions	20 (normalized spatial discrepancy among models 0.74)	>20 calculations yielding similar models.
χ^2 , CorMap P -value	1.12, 0.85	ZBTB38:mCZ38BS, 1.28, 0.124; with mCZ38BS, 0.91, 0.676
Relative phase volumes, R_g values for complex and DNA	n.r.	ZBTB38, 0.519, 26.11; mCZ38BS, 0.482; 20.17
Atomistic modelling/software		<i>SASREF7</i>
q -range for fit (\AA^{-1})		0.008–0.336
Symmetry/anisotropy assumptions		$P1$
No. of individual model reconstructions		>20, best fit selected
χ^2 , CorMap P -value		1.65, 0.03

(f) Data and model deposition.

SASBDB ID	mCZ38BS SASDCB3	ZBTB38:mCZ38BS SASDCA3
-----------	--------------------	---------------------------

† The difference in significant figures here is due to how they are reported in the R_g module of *PRIMUS/qt*.

error is reasonable. The SAXS sample for the complex was made by concentrating samples of the complex prepared for NMR measurements and assessing the change in volume upon concentration, which would likely contribute to the large discrepancy for the complex.

(iii) A reviewer might ask: why are the data truncated at $q = 0.2 \text{\AA}^{-1}$ for *DAMMIN* and *MONSA*, but $P(r)$ and *SASREF* modelling is taken to 0.3\AA^{-1} ? The different q -ranges make it difficult to compare χ^2 values; the significance of differences between reduced χ^2 values (*i.e.* with χ normalized to the number of data points) can only be properly assessed for data collected over the same data range with the same data points. There is nothing in the paper to indicate why this choice was made, although one might consider that data beyond $q = 0.2 \text{\AA}^{-1}$ are influenced by the effects of scattering density fluctuations within molecular components, which are not accounted in *DAMMIN* or *MONSA* modelling.

(iv) What is the meaning of the $\chi^2 > 1$ values for the fits to the complex? *Answer:* Given that the CorMap P -values all indicate good fits, a possible explanation is that the errors in the data are systematically underestimated. Downloading the model fits in the SASBDB submission and using the Data Comparison tool in *PRIMUS/qt* reveals relatively flat error-

weighted residual plots for the model versus experiment. The discrepancy may be attributable to an aspect of the data reduction, or perhaps the desmearing procedure for the SAXS modelling.

In this example, we can conclude that the models are reasonable fits to the data, noting that the experimental errors are likely to be underestimated. Further assessment of the atomistic model would require information from the paper on how the components were constructed and the detailed modelling protocol that was used. This example illustrates how tabulated information allows the reader of this paper to quickly consider the key parameters as they review the figures in the paper and the deposited model fits, so that they can draw their own conclusions regarding the quality of the data and the validity of the model with confidence, noting any limitations.

4. Creating the SAS-cv template

The general template table does not include the additional reporting expected for SAS-cv experiments, which are nevertheless described in detail in the 2017 guidelines. There is at least one recent example of a SANS-cv study that followed

the original reporting table template and included many of the recommended additions (Furlong *et al.*, 2018). To provide a template that is applicable to the broader class of SAS-cv experiments, we have created a specific template (Supplementary Table S4). Starting with the updated generic template, additional reporting recommendations have been added as follows.

(i) Information on specific deuteration of components and of the solvent for SANS-cv.

(ii) Information on the concentration/type of contrast agents for SAXS-cv.

(iii) Reporting molecular mass (M) for all contrast points and relevant solvent match points.

(iv) As the calculated Porod volume (V_P) for systems with heterogeneous scattering density does not correlate directly with the particle volume, these values would not generally be reported except in cases where the scattering is from a particle with at least approximately homogeneous scattering density.

(v) A new panel (*f*) *Component structural parameters* is included for a system with two components with differing scattering densities. The results reported here should include Sturhmann and parallel axis analysis, and composite scattering functions. For these systems, the derived composite scattering functions [equivalent to the scattering profiles (I_1 and I_2) of the two components and the cross term (I_{12})] can be summed to give the scattering curve of the protein complex with homogeneous contrast [*i.e.* $I_{\text{homogeneous}}(q) = I_1(q) + I_2(q) + I_{21}(q)$] and V_P for the complex can be determined from this profile. These profiles should be included, along with the measured scattering profiles, in the data deposition set and the parameters derived from them are reported in this section. It is also recommended that the composition scattering functions be represented in a figure in the main text.

4.1. Testing the utility of the SAS-cv template: a two-protein contrast-variation experiment

The SAS-cv template was evaluated for its relative ease of use and utility with data and modelling from a combined SAXS/SANS-cv experiment of a protein complex consisting of a histidine kinase with bound protein inhibitors that were partially deuterated (Whitten *et al.*, 2007; Fig. 1*b*). Table 2 was obtained by populating the template using the data and model deposited in the SASBDB (entry SASDHY3), which was performed several years after the original publication and as a result there are very slight differences in a few parameters compared with those presented in the paper. These differences are attributable to the use of more modern software versions or, in the case of the component scattering functions, updated scattering contrast values.

As for the example above of the SAXS study of the protein–DNA complex, the tabulated data quickly provide the reader/reviewer with a concise, comprehensive view of the samples measured, data-collection details and the results of data analysis and modelling. The reader quickly understands the following.

(i) The target complex, KinA₂–2^DSda, is a histidine kinase dimer with two bound deuterated protein inhibitors of relatively small size.

(ii) The data set includes SAXS data from a pinhole-geometry laboratory-based instrument and seven SANS data sets collected on the NIST 30 m SANS instrument (two detector positions) from samples in 0, 10, 20, 40, 80, 90 and 100%(v/v) D₂O, with three measurements either side of the solvent match point for the complex, and this match point can therefore be determined optimally by interpolation [51%(v/v) D₂O] as opposed to extrapolation.

(iii) There is good agreement between Guinier- and $P(r)$ -derived parameters for SAXS and all of the measurements in the SANS contrast series (for details regarding this criterion, see Section 2.3 of Trehwella *et al.*, 2017).

(iv) Linear correlation coefficients generally indicate good Guinier fits, with the worst fit being for the 40%(v/v) D₂O data (Pearson's R value -0.903). All $P(r)$ model fits to data are excellent based on χ^2 and CorMap P -values.

(v) The experimentally determined M values are all within ~ 10 – 25% of the expected values, and V_P from the SAXS data gives a ratio in the expected range for a hydrated protein.

(vi) Sturhmann and parallel axis analyses, and composite scattering function analysis, each give consistent R_g values for the complex and its components. V_P from $I_{\text{homogeneous}}$ is smaller than that derived from SAXS data, likely due to it being a combination of data that is variously impacted by the contrast of the hydration layer, which depending on the contrast point can be of the same or opposite in sign compared with the protein.

(vii) The χ^2 values for the model fit for each contrast measurement are all near 1, except for 90%(v/v) D₂O SANS (1.15) and SAXS data (1.28), and in the case of the low-contrast/high ¹H measurements are significantly less than 1, most likely due to the overestimation of uncertainties for these measurements. CorMap P -values indicate statistically acceptable fits for all, except for SAXS (P -value 0.001) and 40%(v/v) D₂O SANS (P -value 0.00005).

With this information in hand, the reader will want to interrogate the figures in the paper and potentially the deposited model fits and consider the proposed model as follows.

(i) Was the wavelength and beam-geometry smearing of the SANS data appropriately dealt with? *Answer:* From the paper, it appears that the assumption was that these smearing effects were assumed to be minimal, which at the time of these experiments was a common assumption for smoothly changing SAS profiles from proteins in solution. This assumption is generally reasonable based on a recent benchmarking study of five proteins in solution (Trehwella *et al.*, 2022). A reviewer might have asked either for some discussion or evidence supporting this assumption, or whether the authors had considered explicitly accounting for data smearing as SASREF7 includes the option of a resolution file to smear the theoretical curve for comparison with experiment. [As an aside, there is complication with using a single resolution file for SANS data collected with two detector settings as the

Table 2

SAS contrast-variation experiment details for a histidine kinase (KinA) complexed with deuterated Sda (^DSda).

(a) Sample details.

Organism	<i>Bacillus subtilis</i>	
Source	<i>E. coli</i> expression for KinA and ^D Sda (Whitten <i>et al.</i> , 2007)	
Description of complex	Sporulation kinase A (KinA) dimer with two bound ^D Sda protein inhibitors (KinA ₂ -2 ^D Sda)	
Scattering particle composition	KinA ₂ 2 ^D Sda inhibitors	
Proteins	2 × UniProt ID P16497, amino acids 383–606 with an additional N-terminal GSHM	2 × UniProt ID Q7WY62, amino acids 7–52 with an additional N-terminal GS
Non-exchangeable deuteration (%)	0	86
Sample environment/configuration	SAXS measurements of KinA ₂ -2 ^D Sda SANS measurements of KinA ₂ -2 ^D Sda	
Solvent composition	50 mM Tris, 200 mM NaCl, 150 mM imidazole pH 8.5	
Sample temperature (°C)	20	
In-beam sample cell	15 µl sample in 2 mm capillary	Hellma quartz cylindrical cells 120-QS (outside diameter 22 mm, path length 1.00 mm)
Batch measurements		
KinA ₂ -2 ^D Sda concentration (mg ml ⁻¹) (from A ₂₈₀ , extinction coefficients calculated using <i>ProtParam</i> ; Gasteiger <i>et al.</i> , 2005)	11.9	11.9, 26.6 (40% sample only)

(b) SANS data collection.

Data-acquisition/reduction software	SAXS Bruker software	SANS <i>Igor Pro</i> software (WaveMetrics, Lake Oswego, Oregon, USA) and the SANS macros developed at the NCNR (Kline, 2006)
Source/instrument description	Bruker Nanostar with a HiStar 2D detector, 100 µm pixel size	NG3 30 m SANS (NIST) with Ordella 640 × 640 mm ³ He position-sensitive detector (Glinka <i>et al.</i> , 1998)
Measured <i>q</i> -range(s) (<i>q</i> _{min} – <i>q</i> _{max}) (Å ⁻¹)	~0.017–0.34	<i>q</i> -range 0.03–0.45 and 0.01–0.09 for S-D 1.33 and 5 m, respectively
Method for scaling intensities	Arbitrary units (a.u.)	Absolute (cm ⁻¹) with respect to direct beam intensity
Exposure time(s) (all single exposures)	1 h	Collection times for S-D 5.0 m: 2 h for 0%, 10%, 20%, 40% D ₂ O, 1 h for 80%, 90%, 100% D ₂ O for the 11.9 and 26.6 mg ml ⁻¹ samples. For S-D 1.33 m the SAS measurement times were half of those for S-D 5.0 m.
Additional relevant details		
Wavelength (Å), Δλ/λ (FWHM)	1.5406	5.82, 14.3%
Beam geometry, sample-to-detector distance (S-D)	Pinhole collimation, S-D 0.7 m	Source aperture 50 mm, sample aperture 9.5 mm. S-D 1.33 m (detector offset by 25.00 cm) and 5.00 m (detector centred).

(c) SAS-derived structural parameters.

Methods/software	Guinier (<i>R</i> _g in <i>ATSAS</i> 2.1), <i>P</i> (<i>r</i>) <i>GNOM</i> 4.6							
SAXS or SANS (%D ₂ O)	SAXS	SANS (100)	SANS (90)	SANS (80)	SANS (40)	SANS (20)	SANS (10)	SANS (0)
Guinier analysis								
<i>I</i> (0) ± σ (SAXS, a.u.; SANS, cm ⁻¹)	805 ± 5	0.530 ± 0.002	0.322 ± 0.003	0.1865 ± 0.0024	0.0677 ± 0.0015	0.2204 ± 0.0007	0.364 ± 0.005	0.545 ± 0.006
<i>R</i> _g ± σ (Å)	29.1 ± 0.3	24.7 ± 0.1	24.3 ± 0.3	22.5 ± 0.4	21.6 ± 0.9	29.1 ± 1.0	28.2 ± 0.5	28.3 ± 0.4
<i>qR</i> _g range	0.58–1.28	0.92–1.29	0.98–1.22	0.98–1.32	0.31–1.22	0.90–1.36	0.75–1.32	0.72–1.27
Pearson's <i>R</i> for fit	−0.999	−0.995	−0.994	−0.990	−0.903	−0.946	−0.990	−0.998
<i>P</i> (<i>r</i>) analysis								
<i>I</i> (0) ± σ (SAXS, a.u.; SANS, cm ⁻¹)	800 ± 2	0.5358 ± 0.0014	0.3217 ± 0.0015	0.1895 ± 0.0015	0.0703 ± 0.0014	0.2183 ± 0.0019	0.3665 ± 0.0030	0.5491 ± 0.0027
<i>R</i> _g ± σ (Å)	28.6 ± 0.1	25.1 ± 0.1	24.4 ± 0.1	23.2 ± 0.2	24.2 ± 0.6	28.9 ± 0.2	28.3 ± 0.2	28.7 ± 0.1
<i>d</i> _{max} (Å)	80	75	75	70	70	80	80	80
<i>q</i> -range (Å ⁻¹)	0.02–0.299	0.024–0.479	0.028–0.479	0.032–0.479	0.015–0.480	0.0254–0.479	0.026–0.479	0.025–0.479
Reciprocal-space fit (χ ² , CorMap <i>P</i> -value)	0.95, 0.54	0.67, 0.48	0.43, 0.94	0.59, 0.04	0.55, 0.76	0.71, 0.47	0.98, 0.02	0.70, 0.74

geometrical factors are different. Had a reviewer raised the question, the question of data smearing could have been usefully explored/addressed.]

(ii) Is there an explanation for the relatively poor linear correlation coefficient for the 40%(v/v) D₂O SANS Guinier fit? *Answer:* While not directly addressed in the paper, the answer is yes as the 40%(v/v) D₂O SANS data are the lowest contrast measurement, which means that it will be maximally influenced by small fluctuations in scattering length density within the scattering particle, in this case from the bound deuterated Sda molecules, and density fluctuations within the much larger, approximately solvent-matched KinA₂ will also

contribute. As a result, the appropriate *q*-range for the Guinier approximation to hold is highly uncertain.

(iii) Is the ~10–25% difference in *M* values from *I*(0)/*c* compared with those calculated from the chemical composition acceptable? *Answer:* The paper reports a concentration of 11.9 mg ml⁻¹ for the KinA₂-2^DSda complex, which was then dialysed into the different percentage D₂O solvents, where it is likely that there may have been small changes in concentration that would affect the calculation of *M* from *I*(0)/*c* at the level observed.

(iv) What might the significance for the model be of the relatively high χ² value and low *P*-value for the SAXS fit and

Table 2 (continued)

(d) Scattering particle size and solvent match points.

Methods/software	Solvent match point from linear dependence of $I(0)^{1/2}$ on %D ₂ O. M , partial specific volume (ν) and scattering contrast ($\Delta\rho$) from chemical composition from <i>MULCh</i> (Whitten <i>et al.</i> , 2008), V_P from <i>PRIMUS/qt</i> (ATSAS 3.2; Manalastas-Cantos <i>et al.</i> , 2021), equation (1) in Trehwella <i>et al.</i> (2017) for M from $I(0)/c$.							
	Kin2-2 ^D Sda complex		KinA monomer		Sda monomer			
Solvent match points								
Calculated	53% (ν/ν) D ₂ O		n.a.		n.a.			
Experimental	51% (ν/ν) D ₂ O		n.a.		n.a.			
Partial specific volume ν (cm ³ g ⁻¹)	0.744		0.739		0.745			
Molecular mass M estimates (kDa)								
From chemical composition	62.39 (62.26 for all ¹ H)		25.35		5.584 for all ¹ H, 5.845 for 86% non-exchangeable H deuterated (determined by mass spectrometry)			
Dynamic light scattering (DynaPro)	Reported as 'monodispersed complex'							
M from SAS contrast data	SAXS	SANS (100)	SANS (90)	SANS (80)	SANS (40)	SANS (20)	SANS (10)	SANS (0)
M from $I(0)/c$ (ratio to expected)	75 (1.21)	72 (1.16)	70 (1.12)	78 (1.25)	52 (0.84)	57 (0.93)	57 (0.92)	56 (0.91)
Contrast $\Delta\rho$ (10 ¹⁰ cm ⁻²)	2.707	-2.620	-2.060	-1.500	0.740	1.860	2.420	2.980
V_P (Å ³) (ratio to expected)	94453 (1.52)	—	—	—	—	—	—	—

(e) Modelling.

Method	Rigid-body refinement (<i>SASREF7</i> , <i>ATSAS</i> 2.1; Petoukhov <i>et al.</i> , 2012)							
Symmetry assumptions	Twofold symmetry axis through the centre of the KinA dimerization domain							
No. of repeats	Simulation repeated 14 times to find the best fit with minimal steric clash penalties							
Fit parameters	SAXS	SANS (100)	SANS (90)	SANS (80)	SANS (40)	SANS (20)	SANS (10)	SANS (0)
q -range (Å ⁻¹)	0.02–0.3	0.01–0.3	0.01–0.3	0.01–0.3	0.01–0.3	0.01–0.3	0.01–0.3	0.01–0.3
χ^2	1.28	0.98	1.15	0.93	0.76	0.56	0.63	0.98
CorMap P -value	0.001	0.029	0.02	0.03	5×10^{-5}	0.058	0.015	0.002

(f) Component structural parameters for a two-component scattering density system.

Methods/software	Equations from Olah <i>et al.</i> (1994) as implemented in <i>MULCh</i> (Whitten <i>et al.</i> , 2008). Note: for Stuhmann and parallel axis analyses only, a second set of measurements from lower concentration samples were used (3.7 mg ml ⁻¹ , collected for 3 h for 0% and 10% D ₂ O and 1 h for 20%, 70%, 80%, 90% and 100% D ₂ O).							
V_P from $I_{\text{homogeneous}}(q)$ for the complex (Å ³)	65900 using the method of Fischer <i>et al.</i> (2010)							
Parameters from two-component analysis	R_g , KinA ₂ in complex (Å)	R_g , two bound Sda (Å)	Center of mass separation, KinA ₂ -2Sda (Å)		d_{max} , KinA ₂ (Å)	d_{max} , two Sda (Å)		
Stuhmann plot	25.90 ± 0.10	24.81 ± 0.94	26.43 ± 1.39		n.r.	n.r.		
Parallel axis theorem	26.04 ± 0.10	25.10 ± 0.88	25.60 ± 1.35		n.r.	n.r.		
Composite scattering functions	25.6 ± 0.1	24.9 ± 0.3	n.r.		77	105		

(g) Data and model deposition.

SASBDB deposition ID	Entry SASDHY3 has the SAXS data as the primary entry. The complete contrast-variation experiment (SAXS and SANS data plus modelling results) is made available as additional files in the full-entry zip archive of SASDHY3 as KinA-Sda_SANS.zip
----------------------	--

the low P -value for the 40% (ν/ν) D₂O SANS data? *Answer:* Inspection of the individual model fits provided in the SASBDB deposition using the Data Comparison tool in *PRIMUS/qt* reveals the following.

(1) For the SAXS data, the relatively high χ^2 value in part reflects the relatively small errors in the data, but there is also a small region of misfit in the mid- q region that would be sensitive to domain positions ($q = 0.18\text{--}0.23 \text{ \AA}^{-1}$), suggesting that there is a possibility of improving the model by small adjustments of the positions of components or domains within KinA₂. There also may be a small degree of dynamical shifting among the domains/components that a single structural model cannot completely represent.

(2) For the 40% data, while the χ^2 value is acceptable for the global fit, the CorMap analysis reveals a region of misfit for $q = 0.1\text{--}0.18 \text{ \AA}^{-1}$ that overlaps with the small region of misfit for the SAXS data. The combination of large errors in this lowest contrast data set plus the potential need for small adjustments in the domain positions or perhaps a small change

in conformation of the Sda upon binding could account for the misfit in these data. Another possibility is that the scattering length density assumed in the modelling may need to be adjusted to better reflect the contrast of the particle.

Overall, however, it can be concluded that the model is an excellent fit to the data. Given the number of SAS profiles used in the modelling and combined with information on the quality of the structures of the KinA₂ and Sda components presented in the paper, at the level of understanding the general relationships between the KinA domains and the Sda inhibitor, one can have high confidence. As it turned out, the model was strongly supported by a subsequent crystallographic study of the KinB-Sda complex (Bick *et al.*, 2009).

Thus, we see again that the table aids the reader or reviewer of this paper in quickly, and with confidence, drawing their own conclusions regarding the quality of the data and the validity of the presented model, albeit with some points that if raised by a reviewer and expanded upon by the authors could have benefitted future readers.

5. Conclusions

While much that is in the 2017 guidelines applies generally to biomolecular SAS, and even to SAS generally, it has never been the case that a single template can accommodate every kind of SAS experiment involving a biomolecule. There are important subgroups of biomolecular SAS studies in which the aim is not three-dimensional structure solution (for example biologically relevant nanoparticles, screening experiments, time-resolved studies, mixtures *etc.*) that are vital contributions to the field and where the reporting framework has yet to be defined by those participating groups. Discussions at SAS2022 have led to interest in developing a template for nanoparticle/micelle/bicelle-type structures (Andreas Haahr Larsen, private communications) and such efforts would be most welcome.

Every effort was made to simplify the template tables presented here. Nevertheless, they remain somewhat more complicated, for example compared with the standard Table 1 for crystallography (data-collection and refinement statistics). Unlike in crystallography, for SAS there is no final purification step such as crystallization to ensure that a pure sample is being measured. Further, the measured crystallographic data set includes thousands or even tens of thousands of individual diffraction intensities against which the model is refined and tested. The experimenter using SAS to investigate biomacromolecular structures in solution must first be able to demonstrate that the scattering is from the particle of interest, free of the influence of inter-particle correlations, impurities or aggregates. That done, it is likely that any given one-dimensional experimental SAS profile may be described by more than one three-dimensional model. As a result, much more information is needed to assess data quality, model fits to the data and the questions of uniqueness that frequently depend upon additional experimental information. The task of populating the table is made easier when an experimenter starts out with the template in mind, as substantial parts can be filled out as the experiment and analysis progress. By populating the updated template tables shown here and presenting the recommended figures and additional data from the 2017 guidelines, authors ensure transparency and completeness in their reporting and the broader structural biology community can be increasingly confident in assessing and using the results of biomolecular SAS experiments.

6. Related literature

The following references are cited in the supporting information for this article: Arnold *et al.* (2014), Benecke *et al.* (2014), Bergmann *et al.* (2000), Bressler *et al.* (2015), Brookes & Rocco (2018), Brookes *et al.* (2016), Chen & Hub (2014), Filik *et al.* (2017), Förster *et al.* (2010), Franke *et al.* (2017), Ginsburg *et al.* (2019), Grant (2018), Grishaev *et al.* (2010), Grudin *et al.* (2017), Hajizadeh *et al.* (2018), Hammersley (2016), Hopkins *et al.* (2017), Ilavsky (2012), Ilavsky & Jemian (2009), Knight & Hub (2015), Liu *et al.* (2020), Liu *et al.* (2012), Narayanan *et al.* (2018), Pedersen *et al.* (2013), Perkins *et al.* (2016), Piidov *et al.* (2019), Poitevin *et al.* (2011), Rambo

& Tainer (2013), Schneidman-Duhovny *et al.* (2016) and Sztucki (2021).

Acknowledgements

Open access publishing facilitated by The University of Sydney, as part of the Wiley - The University of Sydney agreement via the Council of Australian University Librarians.

References

- Aalst, W. M. P. van der, Bichler, M. & Heinzl, A. (2017). *Bus. Inf. Syst. Eng.* **59**, 311–313.
- Arnold, O., Bilheux, J. C., Borreguero, J. M., Buts, A., Campbell, S. I., Chapon, L., Doucet, M., Draper, N., Ferraz Leal, R., Gigg, M. A., Lynch, V. E., Markvardsen, A., Mikkelsen, D. J., Mikkelsen, R. L., Miller, R., Palmen, K., Parker, P., Passos, G., Perring, T. G., Peterson, P. F., Ren, S., Reuter, M. A., Savici, A. T., Taylor, J. W., Taylor, R. J., Tolchenov, R., Zhou, W. & Zikovskiy, J. (2014). *Nucl. Instrum. Methods Phys. Res. A*, **764**, 156–166.
- Baker, E. N., Bond, C. S., Garman, E. F., Newman, J., Read, R. J. & van Raaij, M. J. (2022). *IUCrJ*, **9**, 1–2.
- Benecke, G., Wagermaier, W., Li, C., Schwartzkopf, M., Flucke, G., Hoerth, R., Zizak, I., Burghammer, M., Metwalli, E., Müller-Buschbaum, P., Trebbin, M., Förster, S., Paris, O., Roth, S. V. & Fratzl, P. (2014). *J. Appl. Cryst.* **47**, 1797–1803.
- Bergmann, A., Orthaber, D., Scherf, G. & Glatter, O. (2000). *J. Appl. Cryst.* **33**, 869–875.
- Berman, H. M., Adams, P. D., Bonvin, A. A., Burley, S. K., Carragher, B., Chiu, W., DiMaio, F., Ferrin, T. E., Gabanyi, M. J., Goddard, T. D., Griffin, P. R., Haas, J., Hanke, C. A., Hoch, J. C., Hummer, G., Kurisu, G., Lawson, C. L., Leitner, A., Markley, J. L., Meiler, J., Montelione, G. T., Phillips, G. N. Jr, Prisner, T., Rappsilber, J., Schriemer, D. C., Schwede, T., Seidel, C. A. M., Strutzenberg, T. S., Svergun, D. I., Tajkhorshid, E., Trehwella, J., Vallat, B., Velankar, S., Vuister, G. W., Webb, B., Westbrook, J. D., White, K. L. & Sali, A. (2019). *Structure*, **27**, 1745–1759.
- Berman, H. M., Trehwella, J., Vallat, B. & Westbrook, J. D. (2018). *Adv. Exp. Med. Biol.* **1105**, 261–272.
- Berman, H. M., Vallat, B. & Lawson, C. L. (2020). *IUCrJ*, **7**, 630–638.
- Bick, M. J., Lamour, V., Rajashankar, K. R., Gordiyenko, Y., Robinson, C. V. & Darst, S. A. (2009). *J. Mol. Biol.* **386**, 163–177.
- Bressler, I., Kohlbrecher, J. & Thünemann, A. F. (2015). *J. Appl. Cryst.* **48**, 1587–1598.
- Brookes, E. & Rocco, M. (2018). *Eur. Biophys. J.* **47**, 855–864.
- Brookes, E., Vachette, P., Rocco, M. & Pérez, J. (2016). *J. Appl. Cryst.* **49**, 1827–1841.
- Brose, C. A. & Tainer, J. A. (2019). *Curr. Opin. Struct. Biol.* **58**, 197–213.
- Burley, S. K., Kurisu, G., Markley, J. L., Nakamura, H., Velankar, S., Berman, H. M., Sali, A., Schwede, T. & Trehwella, J. (2017). *Structure*, **25**, 1317–1318.
- Chen, P. C. & Hub, J. S. (2014). *Biophys. J.* **107**, 435–447.
- Filik, J., Ashton, A. W., Chang, P. C. Y., Chater, P. A., Day, S. J., Drakopoulos, M., Gerring, M. W., Hart, M. L., Magdysyuk, O. V., Michalik, S., Smith, A., Tang, C. C., Terrill, N. J., Wharmby, M. T. & Wilhelm, H. (2017). *J. Appl. Cryst.* **50**, 959–966.
- Fischer, H., de Oliveira Neto, M., Napolitano, H. B., Polikarpov, I. & Craievich, A. F. (2010). *J. Appl. Cryst.* **43**, 101–109.
- Förster, S., Apostol, L. & Bras, W. (2010). *J. Appl. Cryst.* **43**, 639–646.
- Franke, D., Jeffries, C. M. & Svergun, D. I. (2015). *Nat. Methods*, **12**, 419–422.
- Franke, D., Petoukhov, M. V., Konarev, P. V., Panjkovich, A., Tuukkanen, A., Mertens, H. D. T., Kikhney, A. G., Hajizadeh, N. R., Franklin, J. M., Jeffries, C. M. & Svergun, D. I. (2017). *J. Appl. Cryst.* **50**, 1212–1225.
- Furlong, E. J., Choudhury, H. G., Kurth, F., Duff, A. P., Whitten, A. E. & Martin, J. L. (2018). *J. Biol. Chem.* **293**, 5793–5805.

- Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & Bairoch, A. (2005). *The Proteomics Protocols Handbook*, edited by J. M. Walker, pp. 571–607. New York: Humana Press.
- Ginsburg, A., Ben-Nun, T., Asor, R., Shemesh, A., Fink, L., Tekoah, R., Levartovsky, Y., Khaykelson, D., Dharan, R., Fellig, A. & Raviv, U. (2019). *J. Appl. Cryst.* **52**, 219–242.
- Glinka, C. J., Barker, J. G., Hammouda, B., Krueger, S., Moyer, J. J. & Orts, W. J. (1998). *J. Appl. Cryst.* **31**, 430–445.
- Grant, T. D. (2018). *Nat. Methods*, **15**, 191–193.
- Grishaev, A., Guo, L., Irving, T. & Bax, A. (2010). *J. Am. Chem. Soc.* **132**, 15484–15486.
- Grishaev, A. V. (2017). *Adv. Exp. Med. Biol.* **1009**, 215–227.
- Grudin, S., Garkavenko, M. & Kazennov, A. (2017). *Acta Cryst.* **D73**, 449–464.
- Hajizadeh, N. R., Franke, D., Jeffries, C. M. & Svergun, D. I. (2018). *Sci. Rep.* **8**, 7204.
- Hammersley, A. P. (2016). *J. Appl. Cryst.* **49**, 646–652.
- Helliwell, J. R. (2019). *Struct. Dyn.* **6**, 054306.
- Hopkins, J. B., Gillilan, R. E. & Skou, S. (2017). *J. Appl. Cryst.* **50**, 1545–1553.
- Ilavsky, J. (2012). *J. Appl. Cryst.* **45**, 324–328.
- Ilavsky, J. & Jemian, P. R. (2009). *J. Appl. Cryst.* **42**, 347–353.
- Jacques, D. A., Guss, J. M., Svergun, D. I. & Trehwella, J. (2012). *Acta Cryst.* **D68**, 620–626.
- Jacques, D. A., Guss, J. M. & Trehwella, J. (2012). *BMC Struct. Biol.* **12**, 9.
- Kikhney, A. G., Borges, C. R., Molodenskiy, D. S., Jeffries, C. M. & Svergun, D. I. (2020). *Protein Sci.* **29**, 66–75.
- Kline, S. R. (2006). *J. Appl. Cryst.* **39**, 895–900.
- Knight, C. J. & Hub, J. S. (2015). *Nucleic Acids Res.* **43**, W225–W230.
- Liu, C.-H., Janke, E. M., Li, R., Juhás, P., Gang, O., Talapin, D. V. & Billinge, S. J. L. (2020). *J. Appl. Cryst.* **53**, 699–709.
- Liu, H., Hexemer, A. & Zwart, P. H. (2012). *J. Appl. Cryst.* **45**, 587–593.
- Manalastas-Cantos, K., Konarev, P. V., Hajizadeh, N. R., Kikhney, A. G., Petoukhov, M. V., Molodenskiy, D. S., Panjkovich, A., Mertens, H. D. T., Gruzinov, A., Borges, C., Jeffries, C. M., Svergun, D. I. & Franke, D. (2021). *J. Appl. Cryst.* **54**, 343–355.
- Narayanan, T., Sztucki, M., Van Vaerenbergh, P., Léonardon, J., Gorini, J., Claustre, L., Sever, F., Morse, J. & Boesecke, P. (2018). *J. Appl. Cryst.* **51**, 1511–1524.
- Olah, G. A., Rokop, S. E., Wang, C. L., Blechner, S. L. & Trehwella, J. (1994). *Biochemistry*, **33**, 8233–8239.
- Pedersen, M. C., Arleth, L. & Mortensen, K. (2013). *J. Appl. Cryst.* **46**, 1894–1898.
- Perkins, S. J., Wright, D. W., Zhang, H., Brookes, E. H., Chen, J., Irving, T. C., Krueger, S., Barlow, D. J., Edler, K. J., Scott, D. J., Terrill, N. J., King, S. M., Butler, P. D. & Curtis, J. E. (2016). *J. Appl. Cryst.* **49**, 1861–1875.
- Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V. & Svergun, D. I. (2012). *J. Appl. Cryst.* **45**, 342–350.
- Piiadov, V., Ares de Araújo, E., Oliveira Neto, M., Craievich, A. F. & Polikarpov, I. (2019). *Protein Sci.* **28**, 454–463.
- Poitevin, F., Orland, H., Doniach, S., Koehl, P. & Delarue, M. (2011). *Nucleic Acids Res.* **39**, W184–W189.
- Pozner, A., Hudson, N. O., Trehwella, J., Terroatea, T. W., Miller, S. A. & Buck-Koehntop, B. A. (2018). *J. Mol. Biol.* **430**, 258–271.
- Rambo, R. P. & Tainer, J. A. (2013). *Nature*, **496**, 477–481.
- Sali, A. (2021). *J. Biol. Chem.* **296**, 100743.
- Sali, A., Berman, H. M., Schwede, T., Trehwella, J., Kleywegt, G., Burley, S. K., Markley, J., Nakamura, H., Adams, P., Bonvin, A. M., Chiu, W., Peraro, M. D., Di Maio, F., Ferrin, T. E., Grünwald, K., Gutmanas, A., Henderson, R., Hummer, G., Iwasaki, K., Johnson, G., Lawson, C. L., Meiler, J., Marti-Renom, M. A., Montelione, G. T., Nilges, M., Nussinov, R., Patwardhan, A., Rappsilber, J., Read, R. J., Saibil, H., Schröder, G. F., Schwieters, C. D., Seidel, C. A., Svergun, D., Topf, M., Ulrich, E. L., Velankar, S. & Westbrook, J. D. (2015). *Structure*, **23**, 1156–1167.
- Schneidman-Duhovny, D., Hammel, M., Tainer, J. A. & Sali, A. (2016). *Nucleic Acids Res.* **44**, W424–W429.
- Schneidman-Duhovny, D., Kim, S. J. & Sali, A. (2012). *BMC Struct. Biol.* **12**, 17.
- Schroer, M. A. & Svergun, D. I. (2018). *Emerg. Top. Life Sci.* **2**, 69–79.
- Sztucki, M. (2021). *SAXSutilities2: a Graphical User Interface for Processing and Analysis of Small-Angle X-ray Scattering Data*. <https://zenodo.org/record/5825707>.
- Trehwella, J. (2016). *Curr. Opin. Struct. Biol.* **40**, 1–7.
- Trehwella, J. (2018). *Adv. Exp. Med. Biol.* **1105**, 77–100.
- Trehwella, J. (2022). *Structure*, **30**, 15–23.
- Trehwella, J., Duff, A. P., Durand, D., Gabel, F., Guss, J. M., Hendrickson, W. A., Hura, G. L., Jacques, D. A., Kirby, N. M., Kwan, A. H., Pérez, J., Pollack, L., Ryan, T. M., Sali, A., Schneidman-Duhovny, D., Schwede, T., Svergun, D. I., Sugiyama, M., Tainer, J. A., Vachette, P., Westbrook, J. & Whitten, A. E. (2017). *Acta Cryst.* **D73**, 710–728.
- Trehwella, J., Hendrickson, W. A., Kleywegt, G. J., Sali, A., Sato, M., Schwede, T., Svergun, D. I., Tainer, J. A., Westbrook, J. & Berman, H. M. (2013). *Structure*, **21**, 875–881.
- Trehwella, J., Vachette, P., Bierma, J., Blanchet, C., Brookes, E., Chakravarthy, S., Chatzimagas, L., Cleveland, T. E., Cowieson, N., Crossett, B., Duff, A. P., Franke, D., Gabel, F., Gillilan, R. E., Graewert, M., Grishaev, A., Guss, J. M., Hammel, M., Hopkins, J., Huang, Q., Hub, J. S., Hura, G. L., Irving, T. C., Jeffries, C. M., Jeong, C., Kirby, N., Krueger, S., Martel, A., Matsui, T., Li, N., Pérez, J., Porcar, L., Prangé, T., Rajkovic, I., Rocco, M., Rosenberg, D. J., Ryan, T. M., Seifert, S., Sekiguchi, H., Svergun, D., Teixeira, S., Thureau, A., Weiss, T. M., Whitten, A. E., Wood, K. & Zuo, X. (2022). *Acta Cryst.* **D78**, 1315–1336.
- Valentini, E., Kikhney, A. G., Previtali, G., Jeffries, C. M. & Svergun, D. I. (2015). *Nucleic Acids Res.* **43**, D357–D363.
- Vallat, B., Webb, B., Fayazi, M., Voinea, S., Tangmunarunkit, H., Ganesan, S. J., Lawson, C. L., Westbrook, J. D., Kesselman, C., Sali, A. & Berman, H. M. (2021). *Acta Cryst.* **D77**, 1486–1496.
- Vallat, B., Webb, B., Westbrook, J. D., Sali, A. & Berman, H. M. (2018). *Structure*, **26**, 894–904.
- Whitten, A. E., Cai, S. & Trehwella, J. (2008). *J. Appl. Cryst.* **41**, 222–226.
- Whitten, A. E., Jacques, D. A., Hammouda, B., Hanley, T., King, G. F., Guss, J. M., Trehwella, J. & Langley, D. B. (2007). *J. Mol. Biol.* **368**, 407–420.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., Hoehn, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. & Mons, B. (2016). *Sci. Data*, **3**, 160018.
- Young, J. Y., Westbrook, J. D., Feng, Z., Sala, R., Peisach, E., Oldfield, T. J., Sen, S., Gutmanas, A., Armstrong, D. R., Berrisford, J. M., Chen, L., Chen, M., Di Costanzo, L., Dimitropoulos, D., Gao, G., Ghosh, S., Gore, S., Guranovic, V., Hendrickx, P. M. S., Hudson, B. P., Igarashi, R., Ikegawa, Y., Kobayashi, N., Lawson, C. L., Liang, Y., Mading, S., Mak, L., Mir, M. S., Mukhopadhyay, A., Patwardhan, A., Persikova, I., Rinaldi, L., Sanz-Garcia, E., Sekharan, M. R., Shao, C., Swaminathan, G. J., Tan, L., Ulrich, E. L., van Ginkel, G., Yamashita, R., Yang, H., Zhuravleva, M. A., Quesada, M., Kleywegt, G. J., Berman, H. M., Markley, J. L., Nakamura, H., Velankar, S. & Burley, S. K. (2017). *Structure*, **25**, 536–545.