Operational properties of fluctuation X-ray scattering data

Erik Malmerberg,[a,b] Cheryl A. Kerfeld[a,b,c] and Petrus H. Zwart[a]*

[a]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA, [b]Department of Plant and Microbial Biology, UC Berkeley, Berkeley, California, USA, and [c]DOE Plant Research Laboratory, Michigan State University, East Lansing, Michigan, USA. *Correspondence e-mail: phzwart@lbl.gov

X-ray scattering images collected on timescales shorter than rotation diffusion times using a (partially) coherent beam result in a significant increase in information content in the scattered data. These measurements, named fluctuation X-ray scattering (FXS), are typically performed on an X-ray free-electron laser (XFEL) and can provide fundamental insights into the structure of biological molecules, engineered nanoparticles or energy-related mesoscopic materials beyond what can be obtained with standard X-ray scattering techniques. In order to understand, use and validate experimental FXS data, the availability of basic data characteristics and operational properties is essential, but has been absent up to this point. In this communication, an intuitive view of the nature of FXS data and their properties is provided, the effect of FXS data on the derived structural models is highlighted, and generalizations of the Guinier and Porod laws that can ultimately be used to plan experiments and assess the quality of experimental data are presented.

1. Introduction

In biology, materials science and the energy sciences, structural information provides important insights into the understanding of matter. The link between a structure and its properties can suggest new avenues for designed improvements of materials, nanoparticles and proteins. For samples without long-range order, such as solutions of biological macromolecules, disordered organic polymers or magnetic domains, as well as (partially) ordered materials, such as self-assembled block copolymers, liquid crystals or assemblies of nanoparticles, structural information can be obtained efficiently using traditional small- and wide-angle X-ray scattering (SAXS/WAXS) techniques (Gann et al., 2012; Dyer et al., 2014). Samples lacking long-range order typically display angular isotropic X-ray scattering patterns, where the mean intensity as a function of scattering angle is directly related to the average shape and local organization of the material investigated (Feigin et al., 1987; Glatter & Kratky, 1982).

The isotropic nature of these SAXS/WAXS diffraction patterns is a result of orientational averaging of the scattering species, due to the fact that the timescale of X-ray exposure exceeds that of rotational diffusion. The advent of coherent X-ray sources (Emma et al., 2010; Ishikawa et al., 2012; Vartanyants et al., 2007; Feldhaus et al., 2013; Borland, 2013) such as free-electron lasers (FELs) and ultra-bright synchrotron light sources allows one to reduce the exposure timescale below that of rotational diffusion such that the non-isotropic intensity fluctuations (or speckle) in the scattering pattern can be resolved.
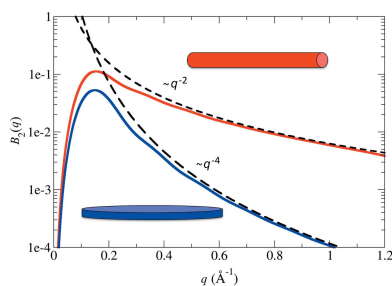
The first experimental demonstration of this technique, termed by the inventor (Kam, 1977) as fluctuation X-ray scattering (FXS), was provided by Kam *et al.* (1981) on frozen tobacco mosaic virus in the early days of synchrotron-based small-angle scattering. Subsequently, fluctuation scattering has been used to detect hidden symmetries in colloids (Wochner *et al.*, 2009) and magnetic domains (Su *et al.*, 2011), for the structure determination of two-dimensional particles (Pedrini *et al.*, 2013; Chen *et al.*, 2012; Saldin, Poon *et al.*, 2010), and for the characterization of liquid crystals (Kurta *et al.*, 2013) and glasses (Cowley, 2001). XFEL-based fluctuation (X-ray) scattering data and structure determination have been demonstrated from single and multiple inorganic nanoparticles (Liu *et al.*, 2013; Mendez *et al.*, 2014) and single polystyrene dumb-bells (Starodub *et al.*, 2012).

Information is extracted from the experimental speckle patterns by computing in-frame angular intensity correlations (Kam, 1977; Saldin, Poon *et al.*, 2010; Saldin, Poon, Bogan *et al.*, 2011; Saldin *et al.*, 2009). These angular intensity correlation curves, the FXS data, can be used for structure determination, either *via* reciprocal-space techniques (Poon *et al.*, 2013; Saldin, Poon, Schwander *et al.*, 2011; Saldin, Shneerson *et al.*, 2010) or *via* real-space methods (Chen *et al.*, 2012; Liu *et al.*, 2013). In earlier studies, FXS has been presented as a method for overcoming experimental and theoretical hurdles in single-particle imaging (Kam, 1977; Saldin *et al.*, 2009). In contrast with this viewpoint, we demonstrate here that FXS is a natural extension of SAXS/WAXS. Despite the increased attention paid to fluctuation scattering due to newly constructed and future light sources, there is a significant lack of understanding of the basic properties of such data. The absence of a basic grasp of the general nature and characteristics of the data makes assessment, validation and proper use of the experimental data a challenge.

This communication will provide an in-depth view of the nature of fluctuation X-ray scattering data, resulting in the derivation of Guinier and Porod relations and other operational properties. We furthermore present the effect of the progressive inclusion of FXS data when reconstructing three-dimensional models, demonstrating the superior quality of models that can be obtained from limited FXS data. The benefits of FXS data apply not only to low-resolution shape or structure determination, but extend to model-based structural refinements as well, allowing one to determine structural changes due to ligand binding or other externally induced perturbations.

2. Results and discussion

2.1. FXS extends traditional small- and wide-angle X-ray scattering

The diffraction pattern of an ensemble of molecules frozen in space and time will contain the signature of many particles, combining effects from the shape and internal structure of the particles, the so-called form factor, and their mutual arrangement in space, the structure factor. In the case of an ideal dilute solution, one can show that the mean angular intensity correlation function, $C_2(q, \Delta\varphi)$, averaged over a large number of independent multiple-particle shots, is equivalent to that obtained from single-particle data (Kam, 1977; Saldin *et al.*, 2009), assuming no interparticle interactions (Kam, 1977; Saldin *et al.*, 2009; Kirian *et al.*, 2011; Altarelli *et al.*, 2010) and the presence of a flat X-ray wavefront during the scattering process (Lehmkühler *et al.*, 2014; Schroer *et al.*, 2014). The potential effects of the coherence properties of the X-ray beam on the resulting angular correlations will be discussed elsewhere.

The angular correlation function can be obtained from the experimental data by averaging a large number of in-frame intensity correlation functions

$$C_2(q, \Delta\varphi) = \left\langle \int_0^{2\pi} I_j(q, \varphi) I_j(q, \varphi + \Delta\varphi) \, d\varphi \right\rangle_{\text{images}}, \quad (1)$$

where $I_j(q, \varphi)$ denotes the intensity as recorded on the j-th diffraction pattern at polar coordinate (q, φ) [$q = (4\pi/\lambda)\sin\theta$, where θ is half the scattering angle and λ is the wavelength of the incident radiation]. Note that additional cross-resolution and n-point correlations can be derived as well (Kam, 1977) but are not considered at this point.

The function $C_2(q, \Delta\varphi)$ can be further decomposed into orthogonal components

$$C_2(q, \Delta\varphi) = \sum_{l=0}^{\infty} B_l(q) F_l(\Delta\varphi), \quad (2)$$

where $B_l(q)$ are resolution-dependent weights and $F_l(\Delta\varphi)$ is given by

$$F_l(\Delta\varphi) = \frac{1}{4\pi} P_l \big[\cos^2\theta(q) + \sin^2\theta(q)\cos(\Delta\varphi) \big]. \quad (3)$$

Here, $P_l(\cdot)$ is a Legendre polynomial and

$$\theta(q) = \frac{\pi}{2} - \sin^{-1}\left(\frac{q}{2\kappa}\right), \quad (4)$$

where κ is equal to the wavenumber $2\pi/\lambda$, with λ the wavelength of the incident radiation. Note that, due to Friedel's law, $B_l(q)$ terms for odd l are equal to 0 (Kam, 1977).

The set of resolution-dependent expansion coefficients $B_l(q)$, as obtained from the experimental data, is related to the three-dimensional structure $\rho(\mathbf{x})$ (Kam, 1977; Saldin *et al.*, 2009). Although the derivation relating the three-dimensional structure to the expansion coefficients $B_l(q)$ is relatively straightforward, it does not provide an intuitive insight into the nature of the data.

Traditionally, fluctuation scattering data are presented starting from the Fourier transform of the real-space structure of the sample (Kam, 1977). Additional insights are obtained when following the route typically used to derive standard relations in small-angle X-ray and neutron scattering. A graphical depiction of fluctuation scattering and how it is related to standard SAXS is shown in Fig. 1, in which the mathematical relations outlined below are referenced. Starting from the real-space structure $\rho(\mathbf{x})$, the Patterson function $\gamma(\mathbf{u})$ can be obtained *via* a self-convolution
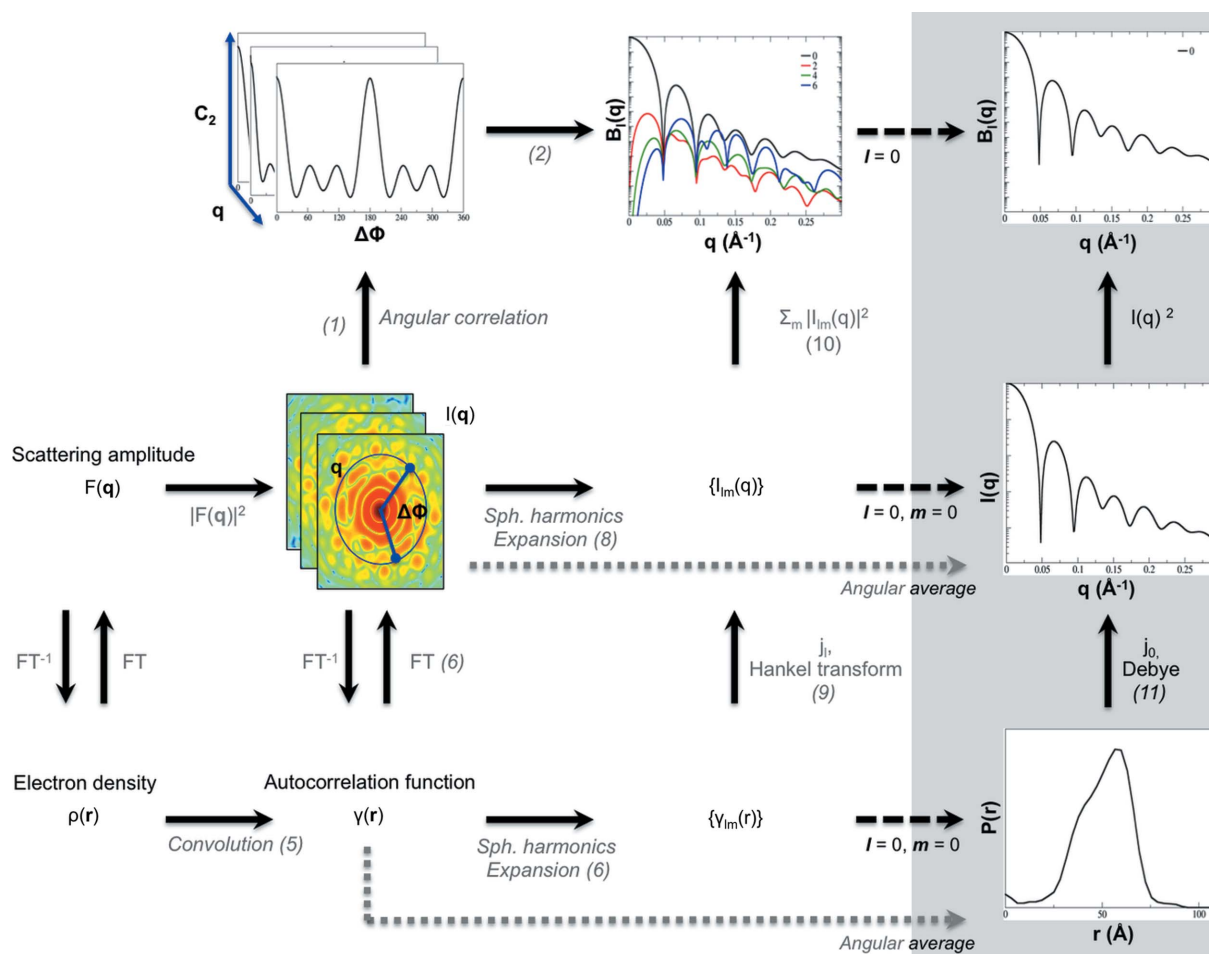
Figure 1
The 'magic square' of scattering is expanded to show the relation between the real-space electron density $\rho(\mathbf{r})$, the associated autocorrelation function $\gamma(\mathbf{r})$ and its Fourier transforms, $F(\mathbf{q})$ and $I(\mathbf{q})$, respectively. When expressing $\gamma(\mathbf{r})$ and $I(\mathbf{q})$ in a spherical coordinate system, Hankel transforms relate the associated expansion coefficients. Orientation-averaged quantities in the grey column, such as SAXS data and the radial distance distribution, can be obtained by selecting curves for which $l = 0$. The numbers in parentheses relate key operations to the corresponding equations given in the text.

$$\gamma(\mathbf{u}) = \int_V \rho(\mathbf{x})\,\rho(\mathbf{x} - \mathbf{u})\,\mathrm{d}\mathbf{x}. \tag{5}$$

By switching to a spherical coordinate system and expressing the Patterson function as a spherical harmonics series, we obtain

$$\gamma(\mathbf{r}) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \gamma_{lm}(r)\,Y_{lm}(\omega_r), \tag{6}$$

where $\gamma_{lm}(r)$ are the expansion coefficient curves of the real-space autocorrelation function and $Y_{lm}(\cdot)$ is a spherical harmonic function. Given that the scattered intensity is proportional to the Fourier transform of the real-space auto-correlation function, one has

$$I(\mathbf{q}) = \int_V \gamma(\mathbf{r})\,\exp(i\mathbf{q}\mathbf{r})\,\mathrm{d}\mathbf{r}. \tag{7}$$

Expressing this intensity function as a spherical harmonics series

$$I(\mathbf{q}) = \sum_{l=0}^{\infty} \sum_{l=-m}^{m} I_{lm}(q)\,Y_{lm}(\omega_q), \tag{8}$$

one obtains (Baddour, 2010)

$$I_{lm}(q) = 4\pi \int_0^{d_{\max}} \gamma_{lm}(r)\,r^2 j_l(qr)\,\mathrm{d}r, \tag{9}$$

where $j_l(\cdot)$ is a spherical Bessel function of order l. These intensity function expansion coefficients are related to the fluctuation scattering curves $B_l(q)$ via (Kam, 1977; Saldin et al., 2009)

$$B_l(q) = \sum_{l=-m}^{m} \left| I_{lm}(q) \right|^2. \tag{10}$$

From the above equations and Fig. 1, it is clear that fluctuation scattering is a natural extension of small-angle X-ray scattering.

In the analysis of traditional SAXS data, the system is assumed to be statistically isotropic, resulting in the assumption that coefficients $I_{lm}(q)$ for $l > 0$ are not experimentally accessible. The term $I_{00}(q)$ is of course equal to SAXS data, as it models the mean intensity as a function of momentum transfer q. Upon further inspection of equation (9) for $l = 0$, we obtain the Debye equation

$$I_{00}(q) = 4\pi \int_0^{d_{max}} \gamma_{00}(r)\, r^2 \, \frac{\sin(qr)}{qr} \, dr, \qquad (11)$$

where $\gamma_{00}(r)r^2$ can be recognized as the pair distance distribution function $P(r)$ (Feigin *et al.*, 1987; Glatter & Kratky, 1982).

Whereas SAXS data only provide experimental information about the zero-order polar Fourier transform of the real-space autocorrelation of the real-space object [equation (11)], fluctuation scattering extends the data into higher-order descriptors of the sample. Given that both SAXS and fluctuation scattering data can be described as *l*-th order spherical Hankel transforms of radial expansion coefficients, it should come as no surprise that certain operational properties from SAXS data can easily be expanded into the fluctuation scattering framework.

3. Guinier and Porod laws for FXS data

As is the case for SAXS data, the low-resolution behaviour of fluctuation scattering data can provide insights into the structural parameters in a model-free fashion and can be used to check the general quality of the data. Using an infinite series expression for spherical Bessel functions (Bowman, 1958) in equation (9) and truncating the series to the second term, as done when deriving the standard Guinier relation, one quickly obtains

$$I_{lm}(q) \approx \hat{I}_{lm} q^l \left(1 - \frac{q^2 R_{lm}^2}{2l+3}\right), \qquad (12)$$

where

$$\hat{I}_{lm} = \frac{Q_{lm}^0}{(2l+1)!!}, \qquad (13)$$

$$R_{lm}^2 = \frac{Q_{lm}^2}{2Q_{lm}^0}. \qquad (14)$$

The quantities Q_{lm}^n are the *n*-th order multipole moments of the autocorrelation function

$$Q_{lm}^n = 4\pi \int_0^{d_{max}} \left[\gamma(\mathbf{r})r^2\right] r^{l+n} \overline{Y}_{lm}(\omega_r)\, d\mathbf{r} \qquad (15)$$

$$= 4\pi \int_0^{d_{max}} P_{lm}(r)\, r^{l+n}\, dr, \qquad (16)$$

with $P_{lm}(r) = \gamma_{lm}(r)r^2$. $\overline{Y}_{lm}(\omega_r)$ denotes complex conjugation of the spherical harmonic $Y_{lm}(\omega_r)$. Note that, in general, $I_{lm}(q)$, $\hat{I}_{lm}$ and R_{lm}^2 are complex quantities unless $l = 0$. Equation (12) can be substituted into equation (10), ultimately resulting in

$$B_l(q) \approx B_l^* \exp\left(2l \log q - \frac{2q^2 R_l^2}{2l+3}\right), \qquad (17)$$

where R_l^2 is equal to the mean real part of R_{lm}^2 $(-l \leq m \leq l)$ and B_l^* is related to the average absolute value of $\hat{I}_{lm}$ for a fixed value of *l*. Linearizing this expression yields a generalized Guinier plot
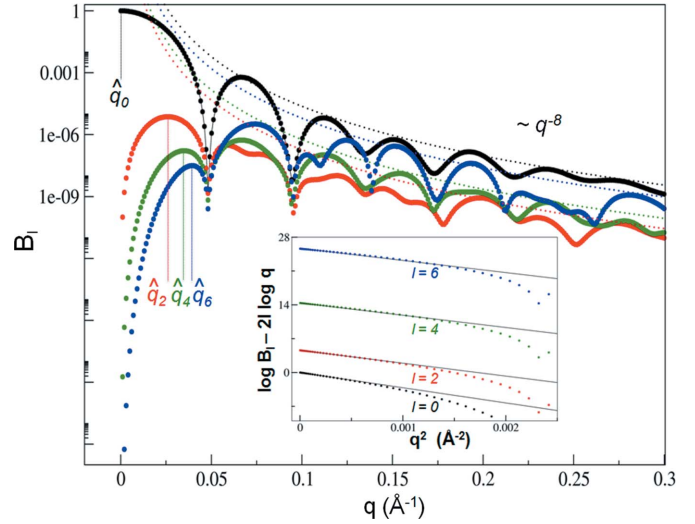

Figure 2
Model $B_l(q)$ coefficients from STMV for $l = 0$ (black), 2 (red), 4 (green) and 6 (blue). The inset depicts generalized Guinier plots with linear fits. The location of the first maximum $\hat{q}_l$ in $B_l(q)$, as obtained from the Guinier analysis, is indicated. The Porod behaviour of the data, characterized by an asymptotic fall-off proportional to q^{-8}, is shown by dotted lines.

$$\left[\log B_l(q) - 2l \log q\right] \approx \log B_l^* - \frac{2q^2 R_l^2}{2l+3}, \qquad (18)$$

where the slope and intercept provide information on the sample-dependent properties B_l^* and R_l^2. From this general formulation of the Guinier equation, it now becomes evident that B_l^* and R_l^2 represent the average amplitudes of the zero- and second-order multipole moments, Q_{lm}^n. For $l = 0$, *i.e.* a monopole, this is synonymous with the square of the total scattering length, $I(0)^2$, and the squared radius of gyration, R_g^2, of the particle. For $l > 0$, these two quantities likewise describe the higher-order moments (quadrupoles, hexadecapoles *etc.*) of the particle shape. The relative magnitudes of these invariants for different values of *l* are influenced by the symmetry of the particle, leading to systematic absences of $B_l(q)$ (Saldin, Poon, Schwander *et al.*, 2011). A generalized Guinier plot from synthetic data is shown in Fig. 2, using satellite tobacco mosaic virus (STMV) as an example.

The generalized Guinier equation also allows one to estimate the location of the first local maximum $\hat{q}_l$ in $B_l(q)$, such that

$$\hat{q}_l = \left[\frac{l(2l+3)}{2}\right]^{1/2} \frac{1}{R_l}, \qquad (19)$$

where the height, $\hat{B}_l$, can be shown to be equal to

$$\hat{B}_l = B_l(\hat{q}_l) = B_l^* \hat{q}_l^{2l} \exp(-l), \qquad (20)$$

Although B_l^* and $\hat{B}_l$ are related, the latter quantity is on a similar numerical scale to the total scattering length, making the use of this quantity more intuitive. $\hat{B}_l$ can be made scale-invariant by normalizing the data such that $B_0(q) = 1$, which is assumed in the following paragraphs.
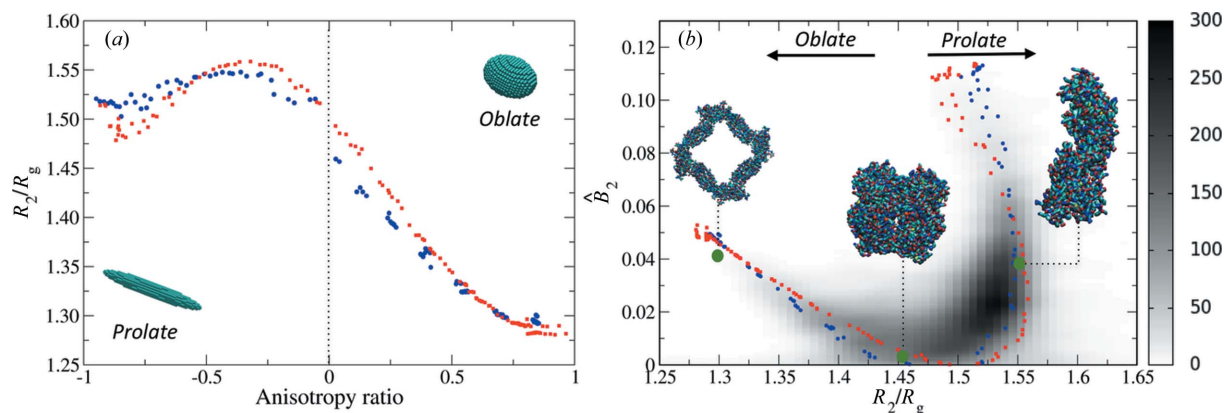
Figure 3
(a) Example shape descriptors for $l = 2$. The ratio R_2/R_g is plotted against the anisotropic ratio for ellipsoids (black dots) or cylinders (red squares), allowing the identification of prolate or oblate features. (b) Including the use of $\hat{B}_l$ as a shape classifier, normalized against $\hat{B}_0$, provides further discriminative power between shapes. Small values of $\hat{B}_2$ represent approximately spherical particles, while large values represent either prolate or oblate particles. The density in part (b) represents the empirical distribution of $(R_2/R_g, \hat{B}_2)$ pairs, as obtained from known PDB structures (see Appendix A for details).

The values $\hat{B}_l$ and R_l can be used as model-free shape classifiers beyond what is provided by the radius of gyration ($R_g = R_0$) as obtained from a standard Guinier analysis. This is exemplified for $l = 2$ in Fig. 3, where the $\hat{B}_2$ values and R_2/R_g ratio have been computed for a number of different sized cylinders, ellipsoids and a representative set of 6709 protein assemblies from the Protein Data Bank (Berman *et al.*, 2000) (see Appendix A for details). From the cylinder and ellipsoid data, it is evident that the combination of R_2/R_g and $\hat{B}_2$ provides a combination of unique shape classifiers that allows one to distinguish prolate from oblate structural features. The value of R_2/R_g indicates whether a shape has prolate or oblate characteristics, while the value of $\hat{B}_2$ measures the extent or strength of the anisotropy, as large values of $\hat{B}_2$ indicate significant deviations from sphericity. Higher-order moments can be used to expand this formalism further to provide a more fine-grained shape classification.

The above generalized Guinier analysis characterizes fluctuation scattering curves at low resolution. For SAXS/WAXS data, high-resolution data trends are described by Porod's law:

$$I(q) \propto q^{-4}. \tag{21}$$

This trend holds for well defined three-dimensional particles. Following Porod's derivation (Feigin *et al.*, 1987; Glatter & Kratky, 1982), but using the l-th order spherical Hankel transform and with an asymptotic approximation for $j_l(\cdot)$ for large q (Bowman, 1958), one readily obtains

$$\left|I_{lm}(q)\right| \propto q^{-4}, \tag{22}$$

and one can thus show that, for large q values, Porod's law extends to fluctuation scattering data

$$B_l(q) \propto q^{-8}. \tag{23}$$

An illustration of this trend for STMV is depicted in Fig. 2.

The Porod behaviour of shapes such as discs [$B_l(q) \propto q^{-4}$] and rods ($B_l \propto q^{-2}$) also displays the same characteristic fall-off (Feigin *et al.*, 1987; Glatter & Kratky, 1982) as expected for squared SAXS intensities (Fig. 4). A practical use of the

predicted Porod behaviour is to use the expected fall-off as an inverse weight when fitting molecular or bead models to data, as is done in SAXS studies (Svergun, 1999).

This combination of Guinier and Porod analyses provides a set of model-independent tools to characterize and validate the quality of the experimental data, in the same way that Guinier and Porod analyses are used in biological small- and wide-angle scattering (Feigin *et al.*, 1987; Glatter & Kratky, 1982). The tools presented here provide straightforward guidelines for the evaluation of experimental FXS data or can be used to plan FXS experiments. An example of the use of the generalized Guinier analysis is the prediction that the first maximum in $B_2(q)$, $\hat{q}_2$, is expected to lie between $2.2/R_g$ and $1.6/R_g$ (see Appendix A). If an R_g estimate is available from standard synchrotron SAXS studies, its value can be used in the experimental design of FEL-based experiments, to ensure that high-quality low-angle FXS data can be obtained.
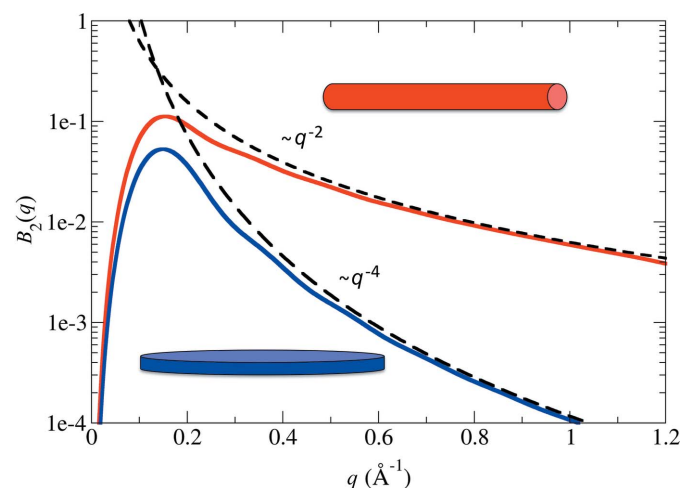
Figure 4
The Porod behaviour of FXS data for one-dimensional rods (red) and two-dimensional discs (blue) follows the same fall-off trends as seen in SAXS/WAXS data. Curves for $l = 2$ are shown; similar trends for higher-order curves exist.
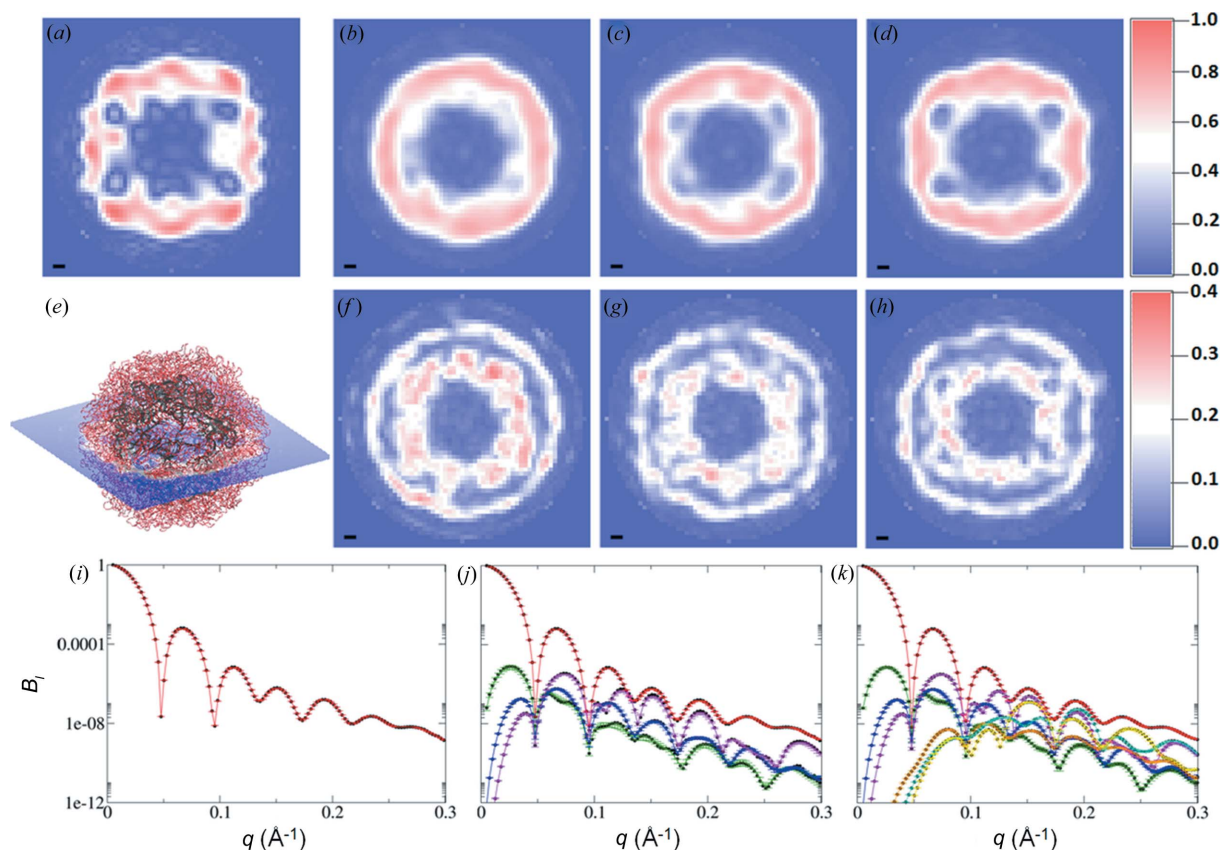
Figure 5
Ab initio unconstrained shape reconstructions from model STMV data. The reconstructions were performed by including $l = 0$ (*i.e.* SAXS data) [parts (*b*), (*f*) and (*i*)], $l \leq 6$ [parts (*c*), (*g*) and (*j*)] or $l \leq 12$ [parts (*d*), (*h*) and (*k*)] in the B_l refinement. The top row [parts (*a*)–(*d*)] depicts density slices through the centre of the virus (*e*). The reference density (*a*) shows significant detail in the core of the virus, which is largely absent when only SAXS data are used (*b*) but which is reproduced, with increasing quality, when terms up to $l = 6$ (*c*) and $l = 12$ (*d*) are considered. Another striking improvement is the distinctly non-spherical outer boundary of the particle when fluctuation scattering data are used. The second row [parts (*f*)–(*h*)] displays the associated standard deviations in the electron density as obtained from the ten independent aligned reconstructions. The black bar [parts (*a*)–(*d*) and (*f*)–(*h*)] represents 10 Å. The bottom row [parts (*i*)–(*k*)] shows the agreement between the data (black circles) and the MOSA-refined (multi-objective simulated annealing; see Appendix A) expansion coefficients [$B_0(q)$ red, $B_2(q)$ green, $B_4(q)$ blue, $B_6(q)$ magenta, $B_8(q)$ orange, $B_{10}(q)$ cyan and $B_{12}(q)$ yellow] for SAXS [part (*i*)] and for fluctuation scattering [parts (*j*)–(*k*)]. The error bars represent the standard deviation from the ten reconstructions.

4. Increased information content

The derivation of the basic properties of FXS data allows one to characterize and evaluate the quality of the experimental data, but fails to explain the reason why these types of experiment are beneficial. The principal advantage of FXS, as shown in Fig. 1, is the additional data made accessible in a fluctuation scattering experiment. This increase in experimental information, even in limited q and l ranges, allows the recovery of more structural detail compared with using $B_0(q)$, *i.e.* the SAXS data alone, in the same q range. This effect is illustrated in Fig. 5, in which average *ab initio* reconstructions obtained from both SAXS and FXS data are shown. The reconstructions are compared with the reference density from which the calculated data were obtained (see Appendix A for details). The reconstructions and analyses here are limited to a relatively low order of l, since these curves are experimentally more easily accessible, and thus provide a conservative overview of the benefit of FXS data compared with standard SAXS data. As is clear from Fig. 5, the addition of limited higher-order scattering information already provides a spectacular increase in reproducible details in the proposed models. One of the reasons why we do not recover the target structure in an error-free fashion is that the optimization problem is still under-constrained (Elser, 2011). However, the main benefit is that FXS is able to reconstruct or derive structural details with greater confidence than can be accomplished from the SAXS data alone, ultimately leading to a better understanding of the structure-related properties.

A similar view of the use of FXS data is obtained when we consider model-based refinement techniques for SAXS/WAXS data (Petoukhov & Svergun, 2005; Gorba & Tama, 2010). Given the stark differences in results obtained in *ab initio* modelling (Fig. 5), the further addition of geometric restraints from a known molecular model could resolve structural ambiguities to such a level that physiologically relevant conformational changes in macromolecules could be confidently deduced from FXS data. For example, when assuming that the structure of a resting state is known, an FXS experiment on the perturbed molecule can provide significantly more data than can be obtained from a SAXS experiment alone. This is illustrated in Fig. 6, in which model FXS
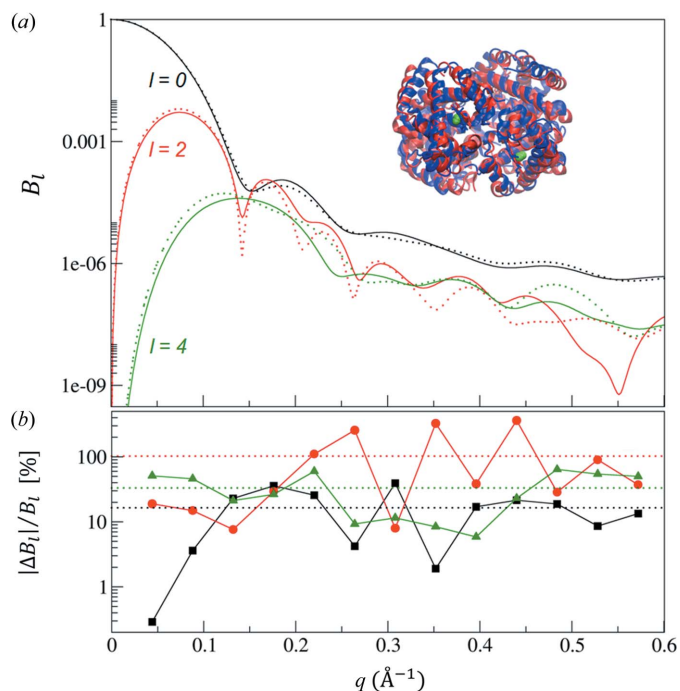
Figure 6
(a) FXS data calculated from the two haemoglobin crystal intermediates 1bbb (CO-haemoglobin; green and red cartoon and dotted lines) and 2hbb (deoxy-haemoglobin; blue cartoon and solid lines) in the Protein Data Bank. The average root mean-square difference between the two intermediates was approximately 2 Å and data were computed for $l \leq 4$ (black, red and green curves). (b) The relative differences, $|\Delta B_l(q)|/B_l(q)$, between the two states at the Shannon sampling points (multiples of $\pi/d_{max} = 0.044$ Å^{-1}) (black squares, red circles and green triangles), indicate the average increased sensitivity of $B_l(q)$ for $l > 0$, as illustrated by the dotted lines. This additional sensitivity, combined with the independent nature of the higher-order curves, ultimately results in a more precise determination of macromolecular structures in solution.

data from carbon monoxide-bound haemoglobin are compared with its unligated intermediate. The relative difference in the data at the Shannon sampling points (Feigin *et al.*, 1987) is depicted as well, indicating that the sensitivity of $B_l(q)$ is enhanced for larger values of *l*. If high-order *l* data up to larger scattering angles are available, difference maps can be obtained as well (Pande *et al.*, 2014). It is worth noting that the extraction of a difference FXS signal will require optimal instrumental and sample conditions, as well as fine-tuned data-processing routines.

This increased information content of FXS data compared with SAXS can play an important role in determining the structural foundation of dynamic processes in biology. As shown earlier (Chen *et al.*, 2013), FXS from a mixture can be described as the component-weighted sum of curves from the individual species. By performing time-resolved FXS experiments, one can obtain $B_l(q)$ curves for intermediate short-lived structural species, akin to standard practices in the analysis of time-resolved WAXS data at synchrotrons (Cammarata *et al.*, 2008; Andersson *et al.*, 2009) or, as recently demonstrated, at an FEL (Arnlund *et al.*, 2014). Thus, the use of fluctuation scattering will ultimately lead to a more accurate depiction of the structural dynamics of macromolecules in solution.

5. Conclusions

In conclusion, we have shown that fluctuation scattering is a natural extension of traditional small-angle X-ray scattering, and that a number of operational properties translate from SAXS/WAXS into fluctuation scattering. Given the increased detail that can be obtained from fluctuation scattering data and the ever-increasing availability of X-ray sources at which these experiments can be performed, we expect that these experiments will become routine in the future. The extended standard Guinier and Porod methods can be used to validate data and characterize samples rapidly in a model-free fashion.

APPENDIX *A*
Additional details

The cylinder and ellipsoid models used in Fig. 3 were obtained by generating voxelized representations of these shapes on a $41 \times 41 \times 41$ voxel cubic grid with 40 Å edges. The set of shapes was obtained by varying their radii and lengths (cylinders) or their main and minor axes (ellipsoids) while keeping a constant volume. The anisotropy ratio as used in Fig. 3 is defined as

$$\frac{\lambda_x - \lambda_z}{\lambda_x + \lambda_z},\tag{24}$$

where λ_z is the moment of inertia, along the axis of revolution for ellipsoids or the cylindrical axis for cylinders. λ_x is the moment of inertia perpendicular to the *z* axis. The anisotropy ratio is 1 for a perfect one-dimensional rod and −1 for a two-dimensional disc. Fig. 3 indicates that R_2/R_g is expected to lie between 1.25 and 1.65. Using equation (19), it follows that the first maximum in $B_2(q)$, $\hat{q}_2$, lies between $2.2/R_g$ and $1.6/R_g$. When determining R_l and $\hat{B}_l$ from FXS curves for $l > 0$, one can either use interpolation and peak picking and equation (19), or use the generalized Guinier transform, equation (18).

The empirical distribution $P(R_2/R_g, \hat{B}_2)$, as shown in Fig. 3(b), was obtained from 6709 PDB files with low (<30%) sequence identity. The distribution shown contains 98% of the density. A small number of structures displayed R_2/R_g ratios below 1.25 or above 1.65, with $\hat{B}_2$ typically close to zero.

All FXS data were computed from either the atomic co-ordinates (PDB models) or the electron density (real-space reconstructions) or voxelized representations (cylinders and ellipsoids) using the three-dimensional Zernike polynomial expansion method (Liu *et al.*, 2012). The maximum expansion order, n_{max}, was determined such that $n_{max} \geq q_{max}r_{max}$, which resulted in $n_{max} = 30$ for STMV and $n_{max} = 40$ for haemoglobin. The $B_l(q)$ coefficients were evaluated to a maximum momentum transfer, q_{max}, of 0.3 Å^{-1} (~20 Å) for STMV (PDB code 1a34) and 0.6 Å^{-1} (~10 Å) for haemoglobin (PDB codes 1bbb and 2hbb). The *ab initio* reconstructions were obtained without the use of symmetry or connectivity restraints *via* a multi-objective simulated annealing (MOSA) (Smith *et al.*, 2008) adaptation of our reverse Monte Carlo procedure (Liu *et al.*, 2013), which has been shown to be less

prone to local minima than when using an aggregated χ^2-based target function. The starting model for the reconstruction was a hollow sphere with a radius of 96 Å on a $61 \times 61 \times 61$ voxel cubic grid.

References

Altarelli, M., Kurta, R. P. & Vartanyants, I. A. (2010). *Phys. Rev. B*, **82**, 104207.

Andersson, M., Malmberg, E., Westenhoff, S., Katona, G., Cammarata, M., Wöhri, A. B., Johansson, L. C., Ewald, F., Eklund, M., Wulff, M., Davidsson, J. & Neutze, R. (2009). *Structure*, **17**, 1265–1275.

Arnlund, D. *et al.* (2014). *Nat Methods*, **11**, 923–926.

Baddour, N. (2010). *J. Opt. Soc. Am. A*, **27**, 2144–2155.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalow, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Borland, M. (2013). *J. Phys. Conference Ser.* **425**, 042016.

Bowman, F. (1958). *Introduction to Bessel Functions*. New York: Dover Publications.

Cammarata, M., Levantino, M., Schotte, F., Anfinrud, P. A., Ewald, F., Choi, J., Cupane, A., Wulff, M. & Ihee, H. (2008). *Nat. Methods*, **5**, 881–886.

Chen, G., Modestino, M. A., Poon, B. K., Schirotzek, A., Marchesini, S., Segalman, R. A., Hexemer, A. & Zwart, P. H. (2012). *J. Synchrotron Rad.* **19**, 695–700.

Chen, G., Zwart, P. H. & Li, D. (2013). *Phys. Rev. Lett.* **110**, 195501.

Cowley, J. M. (2001). *Ultramicroscopy*, **90**, 197–206.

Dyer, K. N., Hammel, M., Rambo, R. P., Tsutakawa, S. E., Rodic, I., Classen, S., Tainer, J. A. & Hura, G. L. (2014). *Methods Mol. Biol.* **1091**, 245–258.

Elser, V. (2011). *Ultramicroscopy*, **111**, 788–792.

Emma, P. *et al.* (2010). *Nat. Photon.* **4**, 641–647.

Feigin, L. A., Svergun, D. I. & Taylor, G. W. (1987). *Structure Analysis by Small-angle X-ray and Neutron Scattering*. New York: Plenum Press.

Feldhaus, J., Krikunova, M., Meyer, M., Möller, T., Moshammer, R., Rudenko, A., Tschentscher, T. & Ullrich, J. (2013). *J. Phys. B*, **46**, 164002.

Gann, E., Young, A. T., Collins, B. A., Yan, H., Nasiatka, J., Padmore, H. A., Ade, H., Hexemer, A. & Wang, C. (2012). *Rev. Sci. Instrum.* **83**, 045110.

Glatter, O. & Kratky, O. (1982). *Small-angle X-ray Scattering*. London: Academic Press.

Gorba, C. & Tama, F. (2010). *Bioinform. Biol. Insights*, **4**, 43–54.

Ishikawa, T. *et al.* (2012). *Nat. Photon.* **6**, 540–544.

Kam, Z. (1977). *Macromolecules*, **10**, 927–934.

Kam, Z., Koch, M. & Bordas, J. (1981). *Proc. Natl Acad. Sci. USA*, **78**, 3559–3562.

Kirian, R. A., Schmidt, K. E., Wang, X., Doak, R. B. & Spence, J. C. H. (2011). *Phys. Rev. E*, **84**, 011921.

Kurta, R. P., Ostrovskii, B. I., Singer, A., Gorobtsov, O. Y., Shabalin, A., Dzhigaev, D., Yefanov, O. M., Zozulya, A. V., Sprung, M. & Vartanyants, I. A. (2013). *Phys. Rev. E*, **88**, 044501.

Lehmkühler, F., Grübel, G. & Gutt, C. (2014). *J. Appl. Cryst.* **47**, 1315–1323.

Liu, H., Poon, B. K., Janssen, A. J. E. M. & Zwart, P. H. (2012). *Acta Cryst.* A**68**, 561–567.

Liu, H., Poon, B. K., Saldin, D. K., Spence, J. C. H. & Zwart, P. H. (2013). *Acta Cryst.* A**69**, 365–373.

Mendez, D., Lane, T. J., Sung, J., Sellberg, J., Levard, C., Watkins, H., Cohen, A. E., Soltis, M., Sutton, S., Spudich, J., Pande, V., Ratner, D. & Doniach, S. (2014). *Philos. Trans. R. Soc. Lond. Ser. B*, **369**, 20130315.

Pande, K., Schwander, P., Schmidt, M. & Saldin, D. K. (2014). *Philos. Trans. R. Soc. Lond. Ser. B*, **369**, 20130332.

Pedrini, B., Menzel, A., Guizar-Sicairos, M., Guzenko, V. A., Gorelick, S., David, C., Patterson, B. D. & Abela, R. (2013). *Nat. Commun*, **4**, 1647.

Petoukhov, M. V. & Svergun, D. I. (2005). *Biophys. J.* **89**, 1237–1250.

Poon, H.-C., Schwander, P., Uddin, M. & Saldin, D. K. (2013). *Phys. Rev. Lett.* **110**, 265505.

Saldin, D. K., Poon, H. C., Bogan, M. J., Marchesini, S., Shapiro, D. A., Kirian, R. A., Weierstall, U. & Spence, J. C. H. (2011). *Phys. Rev. Lett.* **106**, 115501.

Saldin, D. K., Poon, H. C., Schwander, P., Uddin, M. & Schmidt, M. (2011). *Opt. Express*, **19**, 17318–17335.

Saldin, D. K., Poon, H. C., Shneerson, V. L., Howells, M., Chapman, H. N., Kirian, R. A., Schmidt, K. E. & Spence, J. C. H. (2010). *Phys. Rev. B*, **81**, 174105.

Saldin, D. K., Shneerson, V. L., Fung, R. & Ourmazd, A. (2009). *J. Phys. Condens. Matter*, **21**, 134014.

Saldin, D. K., Shneerson, V. L., Howells, M. R., Marchesini, S., Chapman, H. N., Bogan, M., Shapiro, D., Kirian, R. A., Weierstall, U., Schmidt, K. E. & Spence, J. C. H. (2010). *New J. Phys.* **12**, 035014.

Schroer, M. A., Gutt, C. & Grübel, G. (2014). *Phys. Rev. E*, **90**, 012309.

Smith, K., Everson, R., Fieldsend, J., Murphy, C. & Misra, R. (2008). *Evol. Comput. IEEE Trans.* **12**, 323–342.

Starodub, D. *et al.* (2012). *Nat. Commun.* **3**, 1276.

Su, R., Seu, K. A., Parks, D., Kan, J. J., Fullerton, E. E., Roy, S. & Kevan, S. D. (2011). *Phys. Rev. Lett.* **107**, 257204.

Svergun, D. I. (1999). *Biophys. J.* **76**, 2879–86.

Vartanyants, I. A., Robinson, I. K., McNulty, I., David, C., Wochner, P. & Tschentscher, Th. (2007). *J. Synchrotron Rad.* **14**, 453–470.

Wochner, P., Gutt, C., Autenrieth, T., Demmer, T., Bugaev, V., Ortiz, A. D., Duri, A., Zontone, F., Grübel, G. & Dosch, H. (2009). *Proc. Natl. Acad. Sci. USA*, **106**, 11511–11514.