Journal of
Applied
Crystallography
ISSN 0021-8898

Received 21 June 2010
Accepted 11 January 2011

Treatment of hydrogen background in bulk and nanocrystalline neutron total scattering experiments

Katharine Page,[a]* Claire E. White,[a,b] Eben G. Estell,[a] Reinhard B. Neder,[c] Anna Llobet[a] and Thomas Proffen[a]‡

[a]Los Alamos National Laboratory, Lujan Neutron Scattering Center, LANSCE-LC, MS H805, Los Alamos, NM 87545, USA, [b]Department of Chemical and Biomolecular Engineering, University of Melbourne, Victoria 3010, Australia, and [c]Institute for Physics and Condensed Matter, Universität Erlangen-Nürnberg, Staudtstrasse 3, D-91058 Erlangen, Germany. Correspondence e-mail: kpage@lanl.gov

Nuclear incoherent neutron scattering contributions present a challenge in the structural characterization of many classes of materials. This article introduces methods for the correction of nanoparticle, bulk crystalline and amorphous powder neutron scattering data with significant incoherent contributions from hydrogen, and describes the effects the corrections have on the resulting atomic pair distribution function data sets. The approach is presented in the context of the *PDFgetN* data-reduction program [Peterson, Gutmann, Proffen & Billinge (2000). *J. Appl. Cryst.* **33**, 1192].

1. Introduction

Neutron diffraction is a powerful method for probing and characterizing the atomic structure of a material. The analysis of both diffuse and Bragg scattering, the 'total scattering', can yield direct information about the interatomic distances in materials in the form of the pair distribution function (PDF) (Egami & Billinge, 2003). The PDF, $G(r)$, is defined in real space as a function of the pair density, $\rho(r)$,

$$G(r) = 4\pi r[\rho(r) - \rho_0\gamma_0], \tag{1}$$

and in reciprocal space as a sine Fourier transform of the normalized total scattering structure function, $S(Q)$,

$$G(r) = \frac{2}{\pi}\int_0^\infty Q[S(Q) - 1]\sin(Qr)\,\mathrm{d}Q. \tag{2}$$

In these relationships, r is the real-space distance between two atoms, Q is the magnitude of the scattering vector [$Q = (4\pi/\lambda)\sin(\theta/2)$, where θ is the scattering angle and λ is the wavelength of the incident radiation], ρ_0 is the average number density in the sample and γ_0 is a characteristic function for finite-sized objects, often called the nanoparticle form factor (Farrow & Billinge, 2009). $\rho(r)$ gives the probability of finding two atoms i and j separated by a distance $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ in real space, weighted by the scattering lengths b_i and b_j and averaged over all pairs of atoms in the sample (Faber & Ziman, 1965):

$$\rho(r) = \frac{1}{4\pi r^2}\sum_{ij}\frac{b_i b_j}{\langle b\rangle^2}\delta(r - r_{ij}). \tag{3}$$

There are numerous correlation functions commonly used for describing the total scattering function (Keen, 2001). The form shown here is that described by Egami & Billinge (2003), implemented in the data-reduction program *PDFgetN* (Peterson *et al.*, 2000) and the data-analysis software packages *DISCUS* (Proffen & Neder, 1997) and *PDFgui* (Farrow *et al.*, 2007). The construct is convenient for studying crystalline and nanocrystalline materials because the $G(r)$ function scales with r and oscillates around 0 at high r, clearly distinguishing structure up to nanometre length scales.

The quantity measured in a neutron scattering experiment is the double differential cross section per unit solid angle and energy interval (Squires, 1978; Keen, 2001). A number of approximations and corrections are necessary in order to isolate the scattering function $S(Q)$ (containing both diffuse and Bragg scattering) utilized above. This routinely includes a static approximation and corrections for sample-independent background, absorption, multiple scattering and elastic incoherent scattering. The relationships of different correlation functions to elastic, inelastic, coherent and incoherent scattering, and the scattering contributions that warrant empirical subtraction of the corrections described in this paper, are summarized in Fig. 1. The intensity of elastic coherent neutron scattering $I(Q)_{\mathrm{E,coherent}}$ is proportional to the Fourier transform in space of the static pair correlation function $G(r, \omega = 0)$, the probability of finding a particle at position r if there is a particle at $r = 0$ at the same time in the static structure. However, the static scattering function, $S(Q, \omega = 0)$, is not the function commonly accessed experimentally and used in total

‡ Present address: Neutron Scattering Science Division, Oak Ridge National Laboratory, Oak Ridge, TN 37931-6475, USA.

scattering analysis. In practice, data from neutron powder diffraction contain both elastic and inelastic scattering contributions, and the formalism commonly used contains energy-integrated coherent inelastic intensity arising from multi-phonon processes. These contributions are normally ignored, though it has been shown that dynamic information can be extracted for simple systems from the experimental structure factor and experimental PDF by modeling coherent elastic and one-phonon inelastic scattering (Dimitrov *et al.*, 1999). Coherent inelastic scattering is utilized in the recently emerging dynamic PDF method, where the structure factor obtained experimentally is a function of momentum and energy transfer, $S(Q, \omega)$ (McQueeney, 1998; Dmowski *et al.*, 2008). The intensity of this scattering is proportional to the space Fourier transform of the dynamic PDF, $G(r, \omega)$. Note that the static scattering factor, $S(Q, \omega = 0)$, represents a special case of this method.

To use the formalism of the experimental PDF, $G(r)$, the total scattering structure function $S(Q)$ should be isolated. While single-atom self-scattering is subtracted as Laue monotonic diffuse scattering (see *Methods*), the inelastic incoherent scattering contributions, arising from the incoherent cross section of the nuclear scattering from different nuclei in a sample, must be empirically subtracted from the experimental data when present. Part of the intensity of inelastic incoherent scattering (the gray box in Fig. 1) is proportional to the space and time Fourier transforms of the self-correlation function $G_s(r, t)$. This gives correlations between a particle and itself: the probability of finding a particle at position r at a time t when the same particle was at $r = 0$ at $t = 0$. Inelastic incoherent scattering can constitute a significant proportion of the scattering intensity in systems where the incoherent scattering cross section is large compared with the coherent cross section and where energy exchange between the neutron and atom is significant. The case is most apparent for hydrogen-containing materials, including metal–organic frameworks, transition metal hydrides, layered hydroxides, and nanoparticle structures capped or stabilized with hydrogenous ligands. We are far from understanding and modeling the nature of this scattering, even for the simplest of systems (Soper, 2009). These contributions must be subtracted from the measured cross section to recover the contributions that are analyzed in total scattering analysis. All methods for subtraction are empirical in nature and require a number of assumptions and approximations. We describe several of these in the following paragraphs.

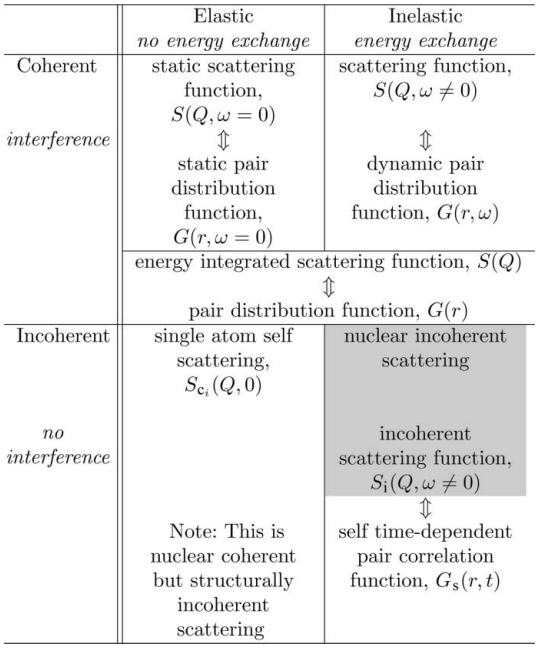

Figure 1
The neutron elastic, inelastic, coherent and incoherent scattering functions, and their real-space counterparts. The experimentally determined pair distribution function (PDF), $G(r)$, must be corrected for incoherent contributions, and contains both elastic and inelastic coherent scattering. The gray box highlights the scattering contributions that warrant empirical subtraction (and the methods of this paper).

When energy exchange is small in comparison with typical system excitations (*i.e.* a liquid), a Placzek correction can be applied. This correction is derived from the ratio of the masses of the neutron and the scattering atom and the first and second moments of $S(Q, \omega)$ (Placzek, 1952) and is not appropriate for solids. Another method applies isotopic substitution to the samples measured. The most common application involves deuterated versions of hydrogen-containing materials (which have a significantly reduced incoherent cross section), and there is a rich body of literature dedicated to this subject. Some success has been found for H-containing samples in measuring a reference sample (*e.g.* water) and subtracting a scale of the molar ratio of H present in the sample and reference data sets ($n\mathrm{H}_{\mathrm{sample}}/n\mathrm{H}_{\mathrm{reference}}$) from the sample scattering (Kameda *et al.*, 2003). The assumption here is that nuclear incoherent hydrogen scattering is similar in the reference and sample data. This may be a poor assumption in the case of extended solids such as transition metal hydrides and H-intercalated compounds. For nanoparticles with hydrogenous ligands and metal–organic framework materials with organic linkers, it may be possible to use as a reference the organic precursors for capping or linking. The authors are unaware of any examples utilizing this approach in solids.

Other approaches are taken in real space. Low-frequency oscillations in Q space (inverse length scales larger than $2\pi/a_0$) only affect distances shorter than the smallest interatomic distance a_0 of the experimental PDF (Peterson *et al.*, 2003). Thus, an experimental $G(r)$ transformed from an uncorrected data set can be truncated below the expected nearest-neighbor pair correlation, extrapolated to zero along an appropriate baseline and back-transformed to $S(Q)$. This particular method is more easily applied to glass or liquid materials where $G(r)$ approaches zero within a few tens of ångströms. Pair correlations are present to very high r for crystalline materials and can cause termination errors in $S(Q)$ during the back Fourier transform. Also, this approach will not remedy inelastic incoherent scattering contributions to the experimental $G(r)$ beyond the first atom–atom distance in real space and is necessarily biased by expected structure results (the nearest-neighbor distance and sometimes the peak shape).

The form of the correction can affect the shape of the experimental PDF and the information held within it, and a systematic exploration of the extent and scale of these effects is called for. We present here a method for correcting hydrogen nuclear incoherent neutron scattering contributions to experimental $S(Q)$ data sets in Q space and examine the effects on the resulting PDF. We use a number of bulk crystalline, nanoparticle and amorphous material total scattering neutron diffraction data sets collected at Los Alamos Neutron Science Center on the NPDF instrument (Proffen *et al.*, 2002) as test cases. Optimization of the process increases the quality of the resulting data, specifically revealing features at low r holding structural information. The present work develops a procedure that can be applied with confidence to experimental data when called for and lays the groundwork for the automation of such a process.

2. Methods

The details of the formalism for the total scattering structure function are rarely shown. Here, we present a partial derivation to accompany our methods and to illustrate the correction procedure using data from a crystalline material.

The complete coherent elastic scattering of N atoms is

$$I(Q)_{\rm E,coherent} = \sum_i^N \sum_j^N b_i b_j \frac{\sin(Qr_{ij})}{Qr_{ij}}. \quad (4)$$

The average intensity, normalized to one atom, is

$$I(Q)_{\rm E,coherent} = \sum_i^N \sum_j^N C_i C_j b_i b_j \frac{\sin(Qr_{ij})}{Qr_{ij}}. \quad (5)$$

In this relationship, b_i is the coherent bound scattering length of atom i averaged over isotopes and spin states, and $C_i = 1/N$. If the self-scattering terms where $i = j$ ($r_{ij} = 0$) are separated, then the scattered-wave interference between distinct ($i \neq j$) atoms in the sample, distinct scattering, can be isolated:

$$I(Q)_{\rm E,coherent} = \sum_i^N \sum_{j,j\neq i}^N C_i C_j b_i b_j \frac{\sin(Qr_{ij})}{Qr_{ij}} + \sum_i^N C_i^2 b_i^2. \quad (6)$$

The average squared b is $\langle b^2\rangle = (1/N)\sum_i^N b^2$. Thus, the equation can be written

$$\begin{aligned} I(Q)_{\rm E,coherent} &= \sum_i^N \sum_{j,j\neq i}^N C_i C_j b_i b_j \frac{\sin(Qr_{ij})}{Qr_{ij}} + \frac{1}{N}\langle b^2\rangle \\ &= \frac{1}{N^2}\sum_i^N \sum_{j,j\neq i}^N b_i b_j \frac{\sin(Qr_{ij})}{Qr_{ij}} + \frac{1}{N}\langle b^2\rangle \\ &= \frac{1}{N}\left[\frac{1}{N}\sum_i^N \sum_{j,j\neq i}^N b_i b_j \frac{\sin(Qr_{ij})}{Qr_{ij}} + \langle b^2\rangle\right] \\ &= \frac{1}{N}\left[NI(Q)_{\rm distinct} + \langle b^2\rangle\right]. \end{aligned} \quad (7)$$

The distinct scattering contains the structural information for the system and is the quantity isolated for total scattering analysis. The self-scattering term or 'single-atom scattering' results from the isotropically averaged scattering intensity for each individual atom, without interference between atoms. It contributes a constant intensity to the measured scattering data.

In the formalism used in the commonly available total scattering programs *DISCUS* (Proffen & Neder, 1997) and *PDFgui* (Farrow *et al.*, 2007), the distinct scattering $I(Q)_{\rm distinct}$ is normalized by the square of the sample average scattering power, $\langle b\rangle^2$, and the total number of atoms, N,

$$S(Q) - 1 = \frac{NI(Q)_{\rm distinct}}{\langle b\rangle^2}. \quad (8)$$

Solving for $NI(Q)_{\rm distinct}$ in equation (7) and substituting yields

$$S(Q) - 1 = \frac{1}{N\langle b\rangle^2}\sum_i^N \sum_j^N b_i b_j \frac{\sin(Qr_{ij})}{Qr_{ij}} - \frac{\langle b^2\rangle}{\langle b\rangle^2}. \quad (9)$$

Rearranged, $S(Q)$ can be defined as

$$\begin{aligned} S(Q) &= \frac{1}{N\langle b\rangle^2}\sum_i^N \sum_j^N b_i b_j \frac{\sin(Qr_{ij})}{Qr_{ij}} - \frac{\langle b^2\rangle}{\langle b\rangle^2} + \frac{\langle b\rangle^2}{\langle b\rangle^2} \\ &= \frac{1}{N\langle b\rangle^2}\sum_i^N \sum_j^N b_i b_j \frac{\sin(Qr_{ij})}{Qr_{ij}} + \frac{\langle b\rangle^2 - \langle b^2\rangle}{\langle b\rangle^2} \\ &= \frac{1}{N\langle b\rangle^2}\sum_i^N \sum_j^N b_i b_j \frac{\sin(Qr_{ij})}{Qr_{ij}} + L, \end{aligned} \quad (10)$$

where $L = (\langle b\rangle^2 - \langle b^2\rangle)/\langle b\rangle^2$ is the normalized Laue monotonic diffuse scattering term, arising from the mixture of chemical species in a sample (nuclear coherent but structurally incoherent scattering). This representation of $S(Q)$ is the one related to $G(r)$ through the Fourier transform in equation (2).

The normalized total scattering structure function above is composed of delta functions that extend throughout all of Q space. The experimentally measured structure function, which we will call $S_{\rm R}(Q)$, is affected by atomic thermal motion, zero-point motion and disorder. These latter effects result in the intensity of the maxima being damped at high Q. The Debye–Waller approximation incorporates lattice vibration effects into crystallographic analysis with a Gaussian function, according to $\exp(-\langle\langle u_i^2\rangle\rangle Q^2)$ (Debye, 1913; Waller, 1923). Here, u_i gives the deviation of atom i from its average position, and the double angle brackets denote averaging over time. As noted earlier, $S_{\rm R}(Q)$ also contains contributions from coherent inelastic scattering from phonons, $S_{\rm I,coherent}$, and may contain significant inelastic incoherent scattering from H, $S_{\rm I,incoherent}$. It has been shown that the total energy-integrated intensity of $S_{\rm I,coherent}$ can be approximated by $1 - \exp(-\langle\langle u^2\rangle\rangle Q^2)$ when summed over large Q (Thorpe *et al.*, 2002; Egami & Billinge, 2003). This gives an expression for the experimentally measured structure function as

$$S_{\rm R}(Q) = \exp\left(-\langle\langle u_i^2\rangle\rangle Q^2\right) S(Q) + S_{\rm I,coherent} + S_{\rm I,incoherent}$$
$$= \exp\left(-\langle\langle u_i^2\rangle\rangle Q^2\right) \frac{1}{N\langle b\rangle^2} \sum_i^N \sum_j^N b_i b_j \frac{\sin\left(Qr_{ij}\right)}{Qr_{ij}}$$
$$+ \exp\left(-\langle\langle u_i^2\rangle\rangle Q^2\right) L + 1 - \exp\left(-\langle\langle u_i^2\rangle\rangle Q^2\right)$$
$$+ S_{\rm I,incoherent}. \quad (11)$$

The goal herein is to separate the contributions of inelastic incoherent scattering from H atoms (which can not be modeled) and produce a corrected total scattering function, $S_{\rm C}(Q)$, that can be used in total scattering analysis:

$$S_{\rm C}(Q) = \exp\left(-\langle\langle u_i^2\rangle\rangle Q^2\right) \frac{1}{N\langle b\rangle^2} \sum_i^N \sum_j^N b_i b_j \frac{\sin\left(Qr_{ij}\right)}{Qr_{ij}}$$
$$+ \exp\left(-\langle\langle u_i^2\rangle\rangle Q^2\right) L + 1 - \exp\left(-\langle\langle u_i^2\rangle\rangle Q^2\right). \quad (12)$$

The first term on the right of equation (12) represents the experimentally measured distinct scattering signal (Bragg and diffuse), which will be defined as $S_{\rm I}(Q)$. The remaining terms form a baseline, $S_{\rm B}(Q)$, for distinct scattering in the absence of inelastic incoherent contributions,

$$S_{\rm B}(Q) = \exp\left(-\langle\langle u_i^2\rangle\rangle Q^2\right) L + 1 - \exp\left(-\langle\langle u_i^2\rangle\rangle Q^2\right). \quad (13)$$

Here again, the value u is defined as the atomic displacement parameter in crystallographic analysis and L contributes the self part of nuclear coherent (but structurally incoherent) scattering. Inspection of equation (13) reveals that this function has a value L at $Q = 0$ and approaches 1 at high Q.

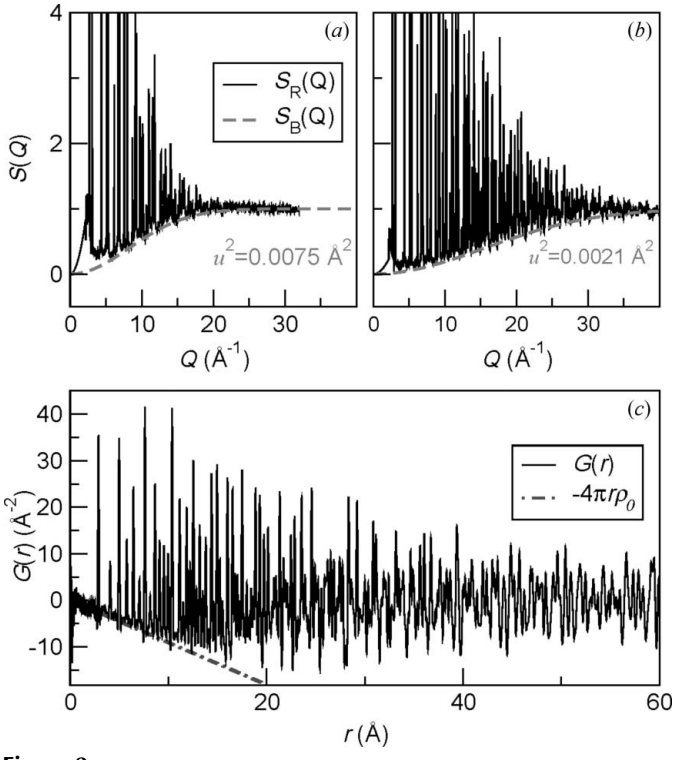

Figure 2
Normalized total scattering patterns (solid lines) and calculated background functions (dashed lines) for a bulk Au sample measured at (*a*) 300 K and (*b*) 15 K. The background function $S_{\rm B}(Q)$ is calculated using refined isotropic displacement parameters for the respective PDF data sets within the Debye–Waller approximation, as defined in equation (13). (*c*) The experimentally determined PDF, $G(r)$, corresponding to the 15 K data in (*b*). The low-r baseline of the PDF, $-4\pi r\rho_0$, is shown as a dashed line.

As a demonstration, Figs. 2(*a*) and 2(*b*) show normalized scattering data from a sample of bulk Au measured at 300 and 15 K, respectively, along with the calculated characteristic background functions $S_{\rm B}(Q)$ for the two temperatures. The total scattering baselines $S_{\rm B}(Q)$ in Figs. 2(*a*) and 2(*b*) were calculated using refined isotropic atomic displacement parameters from a real-space structure refinement of a face-centered cubic (f.c.c.) Au model against the respective PDF data sets. The system displays the behavior expected for a system without an incoherent inelastic signal from H contributions. The PDF, $G(r)$, corresponding to the 15 K data set is shown in Fig. 2(*c*). The PDF follows the $-4\pi r\rho_0$ baseline at low r according to equation (1) ($\gamma_0 = 1$ for a bulk system) and oscillates around zero at high r.

As discussed earlier, nuclear incoherent scattering contributions from hydrogen must be empirically subtracted. The method proposed here involves subtracting all scattering contributions beneath the distinct scattering signal, defined here as $S_{\rm B,exp}(Q)$, and adding back the $S_{\rm B}(Q)$ baseline expected in the absence of nuclear incoherent scattering. It is important to realize that this method assumes that incoherent inelastic scattering contributions are low-frequency components that do not give rise to distinct-like scattering signals. Similar assumptions were made and explained clearly by Soper (2009). A corrected total scattering function (without incoherent inelastic scattering contributions), $S_{\rm C}(Q)$, is then formed as

$$S_{\rm C}(Q) = S_{\rm R}(Q) - S_{\rm B,exp}(Q) + S_{\rm B}(Q) = S_{\rm I}(Q) + S_{\rm B}(Q). \quad (14)$$

We have implemented our correction in the *PDFgetN* data-reduction software (Peterson *et al.*, 2000) as a tool made available in the latest distribution. The procedure and the tool provided may be adapted for any instrument or data set with an $S_{\rm R}(Q)$ input file. The described tool is incorporated for the current configuration of the NPDF instrument, which produces four $S_{\rm R}(Q)$ data sets for four different detector angles. The subtraction must be completed for individual banks of data before the data are merged, as the background scattering intensity from hydrogen is angle dependent. When first processing in *PDFgetN*, automatic scaling of separate banks of data should be turned off and the data should be normalized using the sample density and composition. An intermediate *PDFgetN* file is read by the tool, and for each bank of data background points are located beneath distinct scattering 'peaks'. An interactive plotting tool, displayed in Fig. 3, automatically plots selected points for the background signal [$S_{\rm B,exp}(Q)$], a fit to the points and the resulting $S_{\rm I}(Q)$ for each bank.

To locate background points beneath the distinct scattering peaks, a smooth background polynomial of third order is fitted by standard least-squares refinement to a small Q range of the data. This background polynomial yields an envelope function of the data set within the small Q range by setting the weighting scheme at each data point i to

$$w_i = \exp\left[-K\left(I_{\mathrm{obs},i} - I_{\mathrm{calc},i}\right)\right], \qquad (15)$$

where the value of K is iteratively incremented from $2^0/(I_{\mathrm{obs,max}} - I_{\mathrm{obs,min}})$ to $2^{12}/(I_{\mathrm{obs,max}} - I_{\mathrm{obs,min}})$. Here, $I_{\mathrm{obs,min}}$ and $I_{\mathrm{obs,max}}$ are the minimum and maximum observed intensities within the current Q range. This weighting scheme ensures that all data points with a value of I_{obs} larger than the calculated value I_{calc} are weighted down. The weight decreases exponentially with increasing observed intensity. Thus, the Bragg peaks in the data set are weighted so as not to contribute to the refinement process. Data points with an intensity below the calculated background polynomial, however, are increased exponentially in weight. This weighting scheme pulls the polynomial down such that it will eventually refine to an envelope function below the observed intensities.

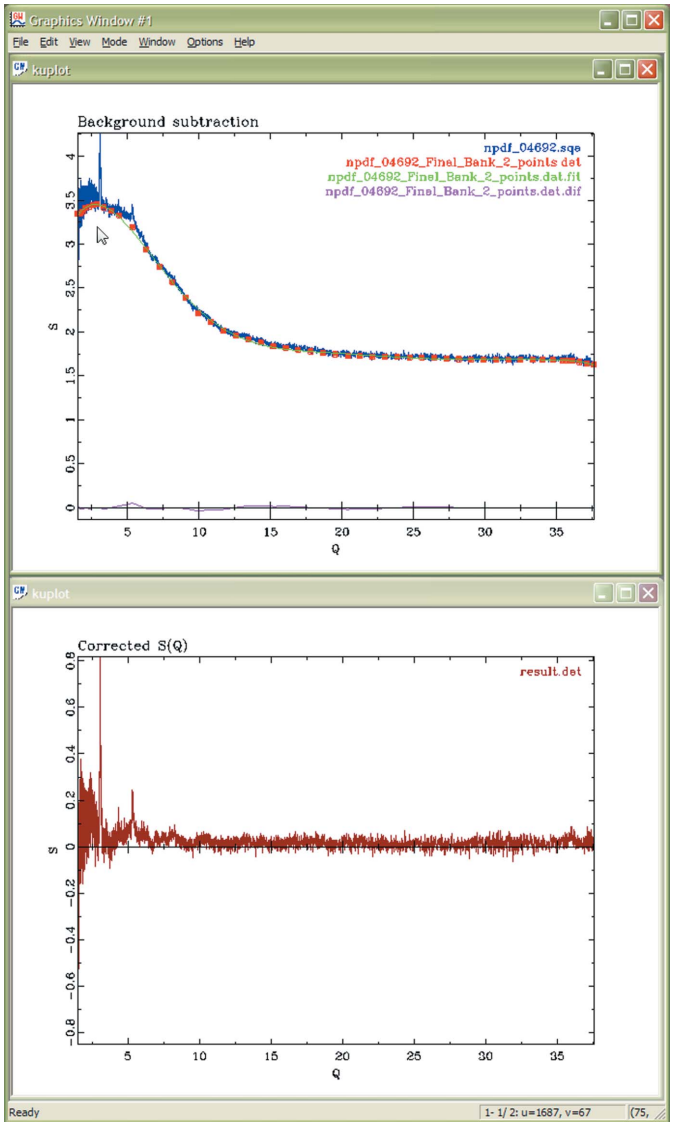

Figure 3
The interactive plotting tool incorporated into *PDFgetN* for subtraction of the experimental background, $S_{\mathrm{B,exp}}(Q)$. $S_{\mathrm{R}}(Q)$ data (blue), background points (red) and a spline interpolation through the background points (green) are displayed in the top panel. The bottom panel displays the resulting $S_{\mathrm{I}}(Q)$, which is automatically updated as points are shifted with a click of the mouse, adjusting the subtracted background.

This process is repeated for all Q ranges of the current detector bank and the midpoints of the calculated function for each interval are taken as interpolation points for the background. The complete background is determined as a spline interpolation between these points. The user can manually move the automatically selected background points with a click of the mouse, adjusting the subtracted $S_{\mathrm{B,exp}}(Q)$ function and resulting $S_{\mathrm{I}}(Q)$. When the user is satisfied with the subtracted background function for each bank of data, the interface is closed and the baseline function $S_{\mathrm{B}}(Q)$ is calculated, utilizing a *PDFgetN* value for L and a user-provided value for $\langle u_i^2 \rangle$. (Depending on the materials of the system, the value of $\langle u_i^2 \rangle$ can be taken from Rietveld refinement, PDF refinement of intermediate data, refinement of a reference sample with no incoherent inelastic contributions, theoretical calculations *etc.* Some of these different strategies are applied in the *Examples* section of this paper.) The $S_{\mathrm{B,exp}}(Q)$ and $S_{\mathrm{B}}(Q)$ functions are combined and saved to a text file, which is uploaded to *PDFgetN* and incorporated prior to merging the banks of data. A corrected total scattering function (without incoherent inelastic scattering contributions), $S_{\mathrm{C}}(Q)$, following equation (14), is then formed.

To illustrate this process and explore the effects of the method on experimentally determined PDF data sets, we systematically examine neutron scattering data from gold nanoparticles capped during seed-mediated growth at approximately 4 nm using a fluorinated octanethiol capping ligand [$CF_3(CF_2)_5(CH_2)_2SH$]. In using a mostly fluorinated ligand, the contributions imparted by inelastic incoherent scattering from hydrogen, while still significant, are minimized. As seen in Fig. 4(*a*), the collected nanoparticle total scattering data, labeled $S_{\mathrm{R}}(Q)$, show broad Bragg-like features atop a significant amount of incoherent scattering centered around 4.5 Å. The PDF resulting from a sine Fourier transform of this data, $G_{\mathrm{R}}(r)$, is given alongside in Fig. 4(*b*). Fig. 4(*c*) shows the intermediate total scattering diffraction pattern $S_{\mathrm{I}}(Q)$, the normalized distinct scattering with inelastic and incoherent scattering subtracted. Fig. 4(*d*) shows $G_{\mathrm{I}}(r)$, which is the PDF resulting from a Fourier transform of the intermediate data. In Fig. 4(*e*) we show our corrected form of the diffraction pattern, $S_{\mathrm{C}}(Q)$. The corresponding PDF, $G_{\mathrm{C}}(r)$, is shown in Fig. 4(*f*).

It is readily apparent that, for the level of nuclear incoherent scattering seen here, subtractions have little effect on the PDF above 2.5 Å in r, conveniently below the nearest-neighbor distance in f.c.c. Au. However, low-r data are strongly affected, with the correction making these correlations follow more closely the expected negative density curve of the function. It is further apparent from comparison of Figs. 4(*d*) and 4(*f*) that inclusion of the background function effects $G(r)$ very little, only changing slightly the intensity of low-r correlations (which correspond to ligand structure in this case). This information is important for two reasons: (*a*) it suggests that, when nuclear incoherent scattering has a low intensity, the PDF is mostly unaffected and we can proceed

with confidence with the corrections if we limit our analysis to the several-ångström range and above (the analysis of f.c.c. Au here); and (*b*) one should take care in correcting and analyzing the low-*r* PDF region (where the correlations of the ligand structure appear in this particular case).

To investigate the statements above more fully, we systematically examined the effects of using different values of $\langle u_i^2 \rangle$ in constructing $S_B(Q)$ and $G_C(r)$. Fig. 5(*a*) gives the gold nanoparticle $S_C(Q)$ at 15 K plotted *versus* $S_B(Q)$ calculated with $u^2 = 0.005$ Å^2 [$S_B^{0.005}(Q)$]. This value of u^2 was taken from refinement of the nanoparticle $G_I(r)$ data. A range of $S_B(Q)$ functions [labeled $S_B^x(Q)$, where $x = u^2$] are displayed along with the data. In Fig. 5(*b*) we compare the effect on the low-*r* $G(r)$ of applying the different $S_B^x(Q)$ functions in the correction procedure. We find that, for $u^2 < 0.01$ Å, the resulting $G(r)$ values are similar above $r = 0.75$ Å, while all of the $G(r)$ values are essentially identical above 2.0 Å. As a check, real-space refinements were completed on all of the reduced data sets shown in Fig. 5(*b*) over the range 2–20 Å in *r* using an f.c.c. Au structure model. A scale factor, lattice parameter and isotropic atomic displacement parameter and peak sharpening were refined in all cases. The refinements gave no differences in structure parameters outside the estimated uncertainty. The applied scale factors for the model were slightly higher for the raw and intermediate PDFs, but identical within error for all the corrected PDF data sets.

3. Examples

The following examples illustrate the value of these techniques for nanoparticle and bulk systems with inelastic incoherent scattering contributions.

3.1. Benzyloxy-capped 5 nm $BaTiO_3$ particles

The current diverse interest in the detailed atomic structure of nanoparticles is manifested in an ever-growing body of total scattering literature dedicated to the topic (for example, Gilbert *et al.*, 2004; Page *et al.*, 2004; Neder & Korsunskiy, 2005; Pradhan *et al.*, 2007; Billinge & Levin, 2007; Masadeh *et al.*, 2007; Chupas *et al.*, 2007). Nanomaterials exhibit lower interference scattering than their bulk counterparts and are often capped or stabilized with hydrogenous ligands, making background subtraction of great importance for the application of neutron total scattering. Here, we show the treatment of data from benzyloxy-capped spherical $BaTiO_3$ nanoparticles (Niederberger *et al.*, 2004). The merged raw data

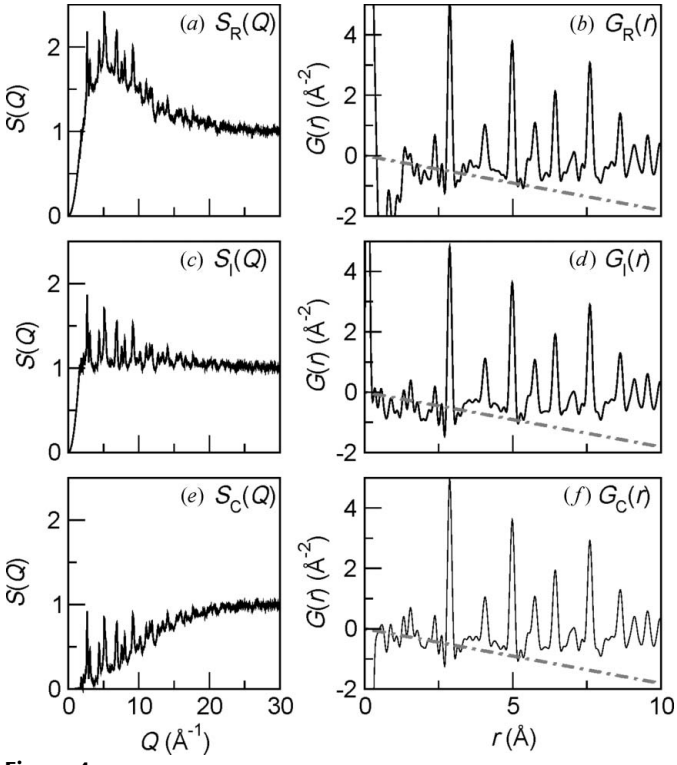

Figure 4
(*a*) The collected nanoparticle diffraction pattern, labeled $S_R(Q)$. (*b*) The PDF $G_R(r)$ resulting from a sine Fourier transform of the data in (*a*). (*c*) The intermediate total scattering diffraction pattern $S_I(Q)$ [normalized total scattering diffraction pattern with background $S_{B,exp}(Q)$ subtracted]. (*d*) $G_I(r)$, which is the PDF resulting from a transform of the intermediate data. (*e*) The corrected form of the diffraction pattern, $S_C(Q)$. The corresponding PDF, $G_C(r)$, is displayed in (*f*). The dot-dashed lines in (*b*), (*d*) and (*f*) show the $-4\pi r\rho_0$ low-*r* baseline of the PDF.

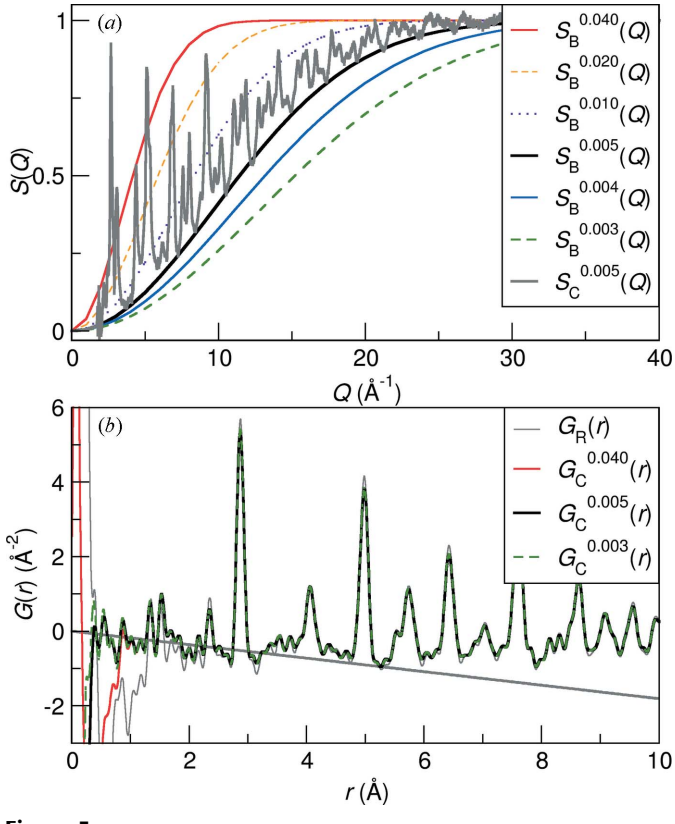

Figure 5
(*a*) Gray line: $S_C(Q)$ for a gold nanoparticle at 15 K plotted for $S_B(Q)$ with $u^2 = 0.005$ Å^2 (value from refinement of the nanoparticle structure). Also displayed are a range of $S_B^x(Q)$ data (see legend), with *x* corresponding to different values of the isotropic atomic displacement parameter u^2 in equation (13). (*b*) The effect on the low-*r* $G(r)$ of applying the different $S_B^x(Q)$ functions from (*a*). For $u^2 < 0.01$ Å, the resulting $G(r)$ values are similar above $r = 0.75$ Å. All the $G(r)$ values are similar above 2.0 Å.

(collected at room temperature from a 0.7 g sample for 9 h), $S_R(Q)$, are displayed along with the merged corrected data, $S_C(Q)$, in Figs. 6(*a*) and 6(*b*). Variations in background scattering intensity with Q arise from nuclear incoherent hydrogen scattering from the benzyloxy capping ligand. In this case, $S_B(Q)$ was derived using U_{iso} parameters from room-temperature bulk $BaTiO_3$ data. The nature of off-centering displacements in these small oxide particles was the subject of the original work presented by Page *et al.* (2010). The careful treatment of background scattering also allowed the structure of the ligands at the surface of the particles to be examined. The results of modeling the low-r region of the corrected PDF, $G_C(Q)$, are shown in Fig. 6(*c*). Contributions to the fit from models of a *P*4*mm* perovskite oxide and a rigid benzyl alcohol molecule are shown below the data and fit.

Phase analysis from refinement gives a benzyl alcohol-to-$BaTiO_3$ mole ratio of 0.35 (3), approximately one capping molecule for every 40 Å^2 of nanoparticle surface. The demonstrated ability to measure quantitatively the oxide particle structure and the surface chemistry of the particle by direct methods is an important advancement in nanoscience characterization. These methods have the potential for further structural interrogation of hybrid organic–inorganic systems, revealing, for example, how ligands bond to surfaces and how they are dispersed over particle surfaces.

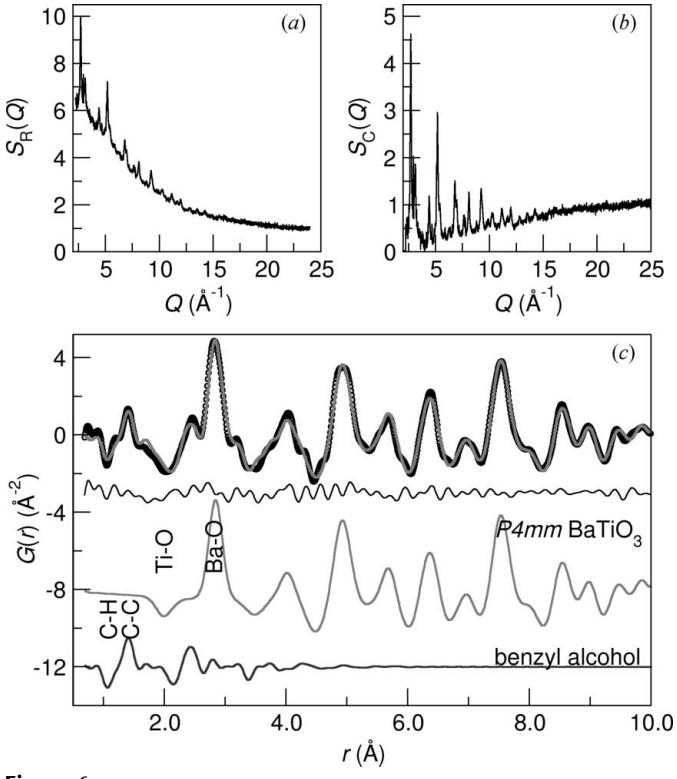

Figure 6
(*a*) The 5 nm $BaTiO_3$ nanoparticle neutron total scattering pattern, labeled $S_R(Q)$, and (*b*) the corrected version, $S_C(Q)$. Incoherent scattering intensity arises from hydrogen in the capping ligand of the particle. (*c*) PDF refinement results for the 5 nm particles. The corrected PDF $G_C(r)$ is displayed as circles, with the fit using a two-phase model for refinement shown as a solid gray line through the data. Below the data and fit are three lines, corresponding to a difference curve for the fit, and the separated 65% phase contribution from $BaTiO_3$ and 35% phase contribution from the benzyloxy ligand. Several near-neighbor peaks are labeled for each phase.

3.2. Kaolinite and its partial dehydroxylate

Kaolinite is a 1:1 phyllosilicate ($Al_2Si_2O_9H_4$) which is used in the production of ceramics, cosmetics, paints and paper, as well as composing a major part of soils worldwide. Upon heating, kaolinite transforms owing to dehydroxylation to a strongly disordered metastable structure, metakaolin (1023 K) (White *et al.*, 2010*a*,*b*). The parent and partially dehydroxylated (773 K) structures contain sufficient hydrogen to require subtraction of background scattering.

The raw scattering pattern for kaolinite is shown in Fig. 7(*a*) as $S_R(Q)$. $G_R(r)$ and $G_C(r)$ are given in Fig. 7(*b*). The value used in the Debye–Waller approximation to give the corrected total scattering pattern $S_C(Q)$ (not shown) was obtained from structure refinement of the intermediate PDF, $G_I(r)$. In Fig. 7(*b*), the raw and corrected scattering patterns show small differences. In particular, the subtraction of the background scattering affects the height of the Si—O, Al—O and O—O peaks (at 1.61, 1.87 and 2.66 Å, respectively), and the differences between the uncorrected and corrected data extend up to 5 Å in r.

$S_R(Q)$ is shown for the partially dehydroxylated sample in Fig. 7(*c*). The material contains less hydrogen than kaolinite and thus displays a smaller degree of background scattering beneath its distinct scattering signal. The same treatment is applied to these data. However, instead of obtaining the average isotropic displacement parameter from a structure refinement (which can be difficult if there is no proposed structure, as is the case for many strongly disordered and amorphous materials), the value is taken from the parent

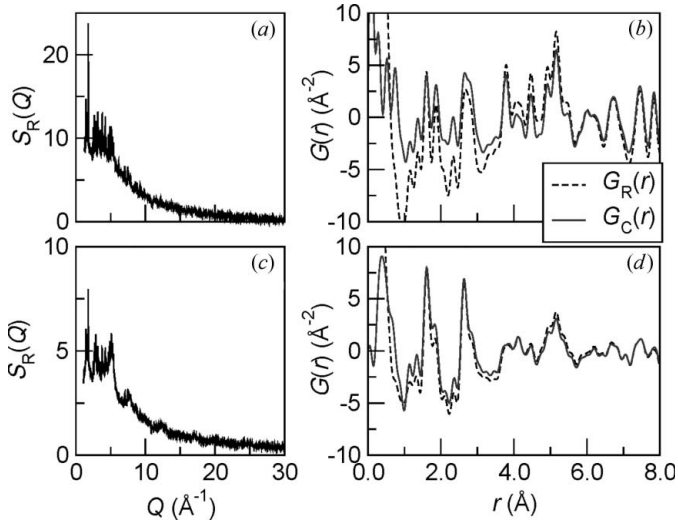

Figure 7
(*a*) The collected kaolinite diffraction pattern, $S_R(Q)$. (*b*) The PDFs $G_R(r)$ and $G_C(r)$, corresponding to a sine Fourier transform of the raw data in (*a*) and the corrected form of the diffraction pattern, $S_C(Q)$, respectively. (*c*) The collected partially dehydroxylated diffraction pattern, $S_R(Q)$. (*d*) The corresponding PDFs $G_R(r)$ and $G_C(r)$ for the data in (*c*).

crystalline material. This choice should be clearly stated in any application of the methods given herein. The corresponding $G_R(r)$ and $G_C(r)$ are displayed in Fig. 7(*d*). Note that the difference between these $G_R(r)$ and $G_C(r)$ is not as large as that in kaolinite, because of the smaller hydrogen component in this case. Only after subtracting the background scattering is it possible to carry out accurate structural comparisons, as has been shown in previous investigations using density functional theory simulations (White *et al.*, 2010*a*).

3.3. Conclusion and outlook

We have presented methods for the removal of hydrogen nuclear incoherent scattering contributions to nanoparticle, bulk crystalline and amorphous powder neutron scattering data. The semi-empirical method produces a corrected total scattering function, which we call $S_C(Q)$. It involves removing all scattering contributions from beneath a normalized distinct scattering signal and incorporating a characteristic baseline, $S_B(Q)$, from calculated or measured quantities to reintroduce the contribution expected from a coherent inelastic signal and nuclear coherent (sample self-) scattering. We have demonstrated the affects of these corrections on the resulting PDF data sets, showing how the low-*r* region is affected to varying degrees and over varying length scales (depending on the intensity of background scattering), while the high-*r* region is unaffected. The proper scale and behavior of the low-*r* region is important for extracting quantitative structural information for systems presenting large amounts of background from nuclear incoherent scattering, as we have shown for both nanoparticle and bulk systems containing hydrogen. The robust nature of the structural refinement of nanoparticle data shows that these methods can be applied with confidence when called for. Step-by-step instructions, a programming tool and example data for these procedures are available in the current *PDFgetN* distribution (Peterson *et al.*, 2000). It is hoped that the methods and quality measures herein will be incorporated into data-reduction software in the future. Finally, we reiterate the importance of clearly stating and understanding the implications of the assumptions made in processing total scattering data in this or any similar manner; indeed, this principle has motivated the present work.

This work has benefited from the use of the NPDF beamline at the Lujan Center at Los Alamos Neutron Science Center, funded by the US DOE, Office of Basic Energy Sciences. Los Alamos National Laboratory is operated by Los Alamos National Security LLC under DOE contract No. DE-AC52-06NA25396. The participation of CEW was funded in part by the Australian Research Council (ARC), including some funding *via* the Particulate Fluids Processing Centre for Sustainable Resource Processing through the Geopolymer Alliance.

References

Billinge, S. J. L. & Levin, I. (2007). *Science*, **316**, 561–565.
Chupas, P. J., Chapman, K. W., Jennings, K. W., Lee, P. L. & Grey, C. P. (2007). *J. Am. Chem. Soc.* **129**, 13822–13824.
Debye, P. (1913). *Verh. Dtsch. Phys. Ges.* **15**, 678.
Dimitrov, D. A., Louca, D. & Röder, H. (1999). *Phys. Rev. B*, **60**, 6204–6207.
Dmowski, W., Vakhrushev, S. B., Jeong, I.-K., Hehlen, M. P., Trouw, F. & Egami, T. (2008). *Phys. Rev. Lett.* **100**, 137602.
Egami, T. & Billinge, S. J. L. (2003). *Underneath the Bragg Peaks: Structural Analysis of Complex Materials*. Amsterdam: Elsevier.
Faber, T. E. & Ziman, J. M. (1965). *Philos. Mag.* **11**, 153–173.
Farrow, C. L. & Billinge, S. J. L. (2009). *Acta Cryst.* A**65**, 232–239.
Farrow, C. L., Juhas, P., Liu, J. W., Bryndin, D., Bozin, E. S., Bloch, J., Proffen, Th. & Billinge, S. J. L. (2007). *J. Phys. Condens. Matter*, **19**, 335219.
Gilbert, B., Huang, F., Zhang, H., Waychunas, G. A. & Bandfield, J. F. (2004). *Science*, **305**, 651–654.
Kameda, Y., Sasaki, M., Usuki, T., Otomo, T., Itoh, K., Suzuya, K. & Fukunaga, T. (2003). *J. Neutron Res.* **11**, 153–163.
Keen, D. A. (2001). *J. Appl. Cryst.* **34**, 172–177.
Masadeh, A. S., Bozin, E. S., Farrow, C. L., Paglia, G., Juhas, P., Billinge, S. J. L., Karkamkar, A. & Kanatzidis, M. G. (2007). *Phys. Rev. B*, **76**, 115413.
McQueeney, R. J. (1998). *Phys. Rev. B*, **57**, 10560–10568.
Neder, R. B. & Korsunskiy, V. I. (2005). *J. Phys. Condens. Matter*, **17**, S125–S134.
Niederberger, M., Garnweitner, G., Pinna, N. & Antonietti, M. (2004). *J. Am. Chem. Soc.* **126**, 9120–9126.
Page, K., Proffen, T., Niederberger, M. & Seshadri, R. (2010). *Chem. Mater.* **22**, 4386–4391.
Page, K., Proffen, T., Terrones, H., Terrones, M., Lee, L., Yang, Y., Stemmer, S., Seshadri, R. & Cheetham, A. K. (2004). *Chem. Phys. Lett.* **393**, 385–388.
Peterson, P. F., Božin, E. S., Proffen, Th. & Billinge, S. J. L. (2003). *J. Appl. Cryst.* **36**, 53–64.
Peterson, P. F., Gutmann, M., Proffen, Th. & Billinge, S. J. L. (2000). *J. Appl. Cryst.* **33**, 1192.
Placzek, G. (1952). *Phys. Rev.* **86**, 377–388.
Pradhan, S. K., Mao, Y., Wong, S. S., Chupas, P. & Petkov, V. (2007). *Chem. Mater.* **19**, 6180–6186.
Proffen, T., Egami, T., Billinge, S. J. L., Cheetham, A. K., Louca, D. & Parise, J. B. (2002). *Appl. Phys. A*, **74**, S163–S165.
Proffen, Th. & Neder, R. B. (1997). *J. Appl. Cryst.* **30**, 171–175.
Soper, A. K. (2009). *Mol. Phys.* **107**, 1667–1684.
Squires, G. L. (1978). *Introduction to the Theory of Thermal Neutron Scattering*. Cambridge University Press.
Thorpe, M. F., Levashov, V. A., Lei, M. & Billinge, S. J. L. (2002). *From Semiconductors to Proteins: Beyond the Average Structure*, edited by S. J. L. Billinge & M. F. Thorpe, pp. 105–128. New York: Kluwer, Plenum.
Waller, I. (1923). *Z. Phys.* **17**, 398–408.
White, C. E., Provis, J. L., Proffen, T., Riley, D. P. & van Deventer, J. S. J. (2010*a*). *Phys. Chem. Chem. Phys.* **12**, 3239–3245.
White, C. E., Provis, J. L., Proffen, T., Riley, D. P. & van Deventer, J. S. J. (2010*b*). *J. Phys. Chem. A*, **114**, 4988–4996.