

Toward a quantitative description of solvation structure: a framework for differential solution scattering measurements

Niklas B. Thompson,* Karen L. Mulfort and David M. Tiede*

Division of Chemical Sciences and Engineering, Argonne National Laboratory, 9700 South Cass Avenue, Lemont, Illinois 60439 USA. *Correspondence e-mail: niklas@anl.gov, tiede@anl.gov

Received 17 January 2024

Accepted 15 April 2024

Edited by I. Robinson, UCL, United Kingdom

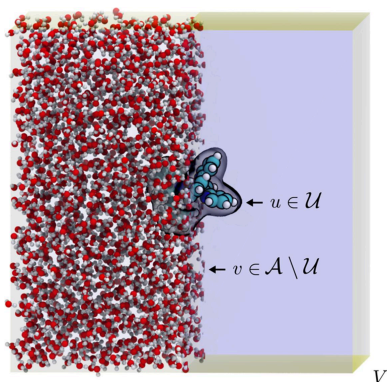
Keywords: aperiodic structures; high-energy X-ray scattering; inorganic chemistry; intermolecular interactions; pair distribution function analysis; SAXS; WAXS; HEXS.

Supporting information: this article has supporting information at www.iucrj.org

Appreciating that the role of the solute–solvent and other outer-sphere interactions is essential for understanding chemistry and chemical dynamics in solution, experimental approaches are needed to address the structural consequences of these interactions, complementing condensed-matter simulations and coarse-grained theories. High-energy X-ray scattering (HEXS) combined with pair distribution function analysis presents the opportunity to probe these structures directly and to develop quantitative, atomistic models of molecular systems *in situ* in the solution phase. However, at concentrations relevant to solution-phase chemistry, the total scattering signal is dominated by the bulk solvent, prompting researchers to adopt a differential approach to eliminate this unwanted background. Though similar approaches are well established in quantitative structural studies of macromolecules in solution by small- and wide-angle X-ray scattering (SAXS/WAXS), analogous studies in the HEXS regime—where sub-ångström spatial resolution is achieved—remain underdeveloped, in part due to the lack of a rigorous theoretical description of the experiment. To address this, herein we develop a framework for differential solution scattering experiments conducted at high energies, which includes concepts of the solvent-excluded volume introduced to describe SAXS/WAXS data, as well as concepts from the time-resolved X-ray scattering community. Our theory is supported by numerical simulations and experiment and paves the way for establishing quantitative methods to determine the atomic structures of small molecules in solution with resolution approaching that of crystallography.

1. Introduction

Solvation is a ubiquitous phenomenon that profoundly affects the physiochemical properties of molecular systems. Underlying these effects is the ‘solvation structure’, including how dissolution affects the atomic structure of the solute as well as the way in which the solvent is ordered locally about the solute. Even in the simple case of outer-sphere self-exchange, the organization of solvent about the exchanging ions makes a decisive contribution to the rate of electron transfer, as eloquently expounded by the work of Marcus (1993). Despite the long history of Marcus theory, it is only with the recent advent of the X-ray free electron laser (XFEL) that experimentalists have succeeded in detecting the structural rearrangement of the solute–solvent interface at the ultrafast timescale of photoinduced electron dynamics (van Driel *et al.*, 2016; Khakhulin *et al.*, 2019; Biasin *et al.*, 2021; Katayama *et al.*, 2023). More broadly, synthetic chemists have long sought to engineer the outer coordination spheres of transition metal complexes to engender favorable solvation interactions in catalyst and device design, often taking inspiration from natural metalloenzymes (Drover, 2022; Ghosh *et al.*, 2022).



Likewise, manipulating the solvation structure of cations in electrolyte is an area of emerging importance in battery research (Rajput *et al.*, 2018; Cheng *et al.*, 2022).

Although controlling solvation structure and dynamics remains a key objective in chemical research, experimental approaches to resolve such structures at the atomic scale are underdeveloped. Spectroscopic techniques such as nuclear magnetic resonance (Clare & Gronenborn, 1991; Bifulco *et al.*, 2007), electron paramagnetic resonance (Attanasio, 1986; Mustafi & Makinen, 1988), vibrational (Perera *et al.*, 2008; Sun & Petersen, 2017) or X-ray absorption spectroscopy (Chen, 2004; Penfold *et al.*, 2014; Jay *et al.*, 2020) can provide valuable insights into the effects of solvation, but remain indirect measures of structure. The extended X-ray absorption fine structure region of *K*-edge X-ray absorption spectroscopy data—a type of scattering phenomenon—can provide direct structural information regarding molecules in liquid or frozen solution; however, typically only the first coordination shell of the absorbing atom can be resolved (Filipponi, 1994; Garino *et al.*, 2014).

In contrast, total scattering measurements provide two-body information out to arbitrary distances with a signal that, in principle, decays only as a result of the actual disorder in the system (Egami & Billinge, 2012; Garino *et al.*, 2014), and are thus uniquely suited to address solvation structure (Hertz, 1970; Soderholm *et al.*, 2005; Megyes *et al.*, 2008; Mulfort *et al.*, 2013; Scalambra *et al.*, 2021). Combining large two-dimensional area detectors with the high-energy incident radiation available at third-generation synchrotron light sources ($h\nu > 50$ keV or $\lambda < 0.25$ Å), it is now possible to routinely collect total X-ray scattering measurements out to large momentum transfers, $Q_{\max} \geq 20$ Å⁻¹, where $Q = 4\pi\sin(\theta/\lambda)$ for photons scattered elastically at an angle of 2θ (Billinge, 2019). Given the duality between reciprocal space (Q) and real space, such measurements yield structural information with sub-Ångström spatial resolution, as is typical in crystallographic analysis. However, while crystallography relies on the presence of Bragg diffraction from a crystalline sample, it is always possible to analyze the total scattered intensity from samples in arbitrary phases of matter.

In the context of total scattering experiments on solutions, where the objective is to determine the solvation structure of, for example, a small molecule, a practical limitation arises from the fact that the scattering signal from the bulk solution dominates the signal from the solute, the latter being proportional to its volume fraction in solution. To address this, a differential approach is typically adopted. That is, two measurements are made, the first of the solution itself, and the second of an otherwise identical sample lacking the solute of interest. Then, adopting the practice employed in small-angle/wide-angle X-ray scattering (SAXS/WAXS) studies of macromolecules (Jeffries *et al.*, 2021), the second measurement is treated as an experimental background and subtracted from the first, with the aim of isolating the solute signal from the bulk solvent signal. Although a theory for subsequent quantitative, atomistic structural modeling has been investigated in the case of SAXS/WAXS data (Park *et al.*, 2009;

Köfinger & Hummer, 2013; Chen & Hub, 2014), where $Q_{\max} \lesssim 1$ Å⁻¹, such treatment is lacking for the case of total scattering data collected in the high-energy regime.

Indeed, unlike in the SAXS/WAXS setting, where spatial resolution is limited, high-energy X-ray scattering (HEXS) measurements are typically transformed to a pair distribution function (PDF) to extract quantitative structural information in real space. Quantitative information (*e.g.* electron densities, bond length distributions *etc.*) can be measured directly from the transformed data when a single peak or a multiplet can be reliably assigned to a specific correlation (Neuefeind *et al.*, 2004; Soderholm *et al.*, 2005; Skanthakumar *et al.*, 2017). However, experimental PDFs often present a complex convolution of many signals, which can be difficult to assign *a priori* without a detailed and realistic structural model for comparison. Nevertheless, we have employed differential HEXS measurements to assign the solution structures of molecular electrocatalysts, generated *in situ* (Yang *et al.*, 2016), and to characterize the atomic scale differences of the intramolecular structures of transition metal complexes between the solid and solution phase (Xie *et al.*, 2022). These studies relied on quantum-mechanical models [*i.e.* density functional theory (DFT) calculations] to predict and assign structural features in the experimental PDFs, with the limitation that any intermolecular (*e.g.* solute–solvent) correlations are neglected.

As we will show, the development of quantitative models of such data requires added complexity, including important contributions to the differential scattering arising from solute–solvent and solvent–solvent pair correlations. The former have been investigated in detail in the context of time-resolved X-ray solution scattering (Kim *et al.*, 2009; Ihee, 2009; Haldrup *et al.*, 2010); however, the principal solvent–solvent correlations differ in origin between the static and time-resolved settings (see below). In this contribution, we present a theoretical framework for the static differential solution scattering experiment from first principles within kinematic scattering theory. Our aim is twofold: first, to clearly delineate the information content contained in the differential scattering intensity and second, to develop a formalism analogous to standard PDF analysis that is applicable to quantitative structural studies of small molecules in solution.

2. Theory

2.1. Standard total scattering formalism

From a structural perspective, the information content contained in the total scattering measurement is encoded by the coherent interference of photons scattered elastically from distinct pairs of atoms within the sample, the coherence volume of the X-rays being typically microscopic (Grübel & Zontone, 2004). The standard formalism for total scattering structural studies involves the Fourier sine transform pair $F(Q) \overset{f}{\leftrightarrow} G(r)$, referred to as the ‘reduced total scattering structure function’ and the ‘reduced PDF’, respectively (Egami & Billinge, 2012).

Consider the scattered intensity from a collection of $\{a_i\}_{i=1,\dots,N} = \mathcal{A}$ atoms within the coherent scattering volume V of the incident X-rays. Within kinematic scattering theory under the independent atom approximation, it is usual to define

$$\langle f(Q)^2 \rangle := \frac{1}{N} \sum_{a_i \in \mathcal{A}} f_i^*(Q) f_i(Q) \quad (1)$$

and

$$\langle f(Q) \rangle^2 := \frac{1}{N^2} \sum_{a_i, a_j \in \mathcal{A}} f_j^*(Q) f_i(Q), \quad (2)$$

where $f_i(Q)$ is the (in principle, complex-valued) atomic form factor of atom a_i (*i.e.* the Fourier transform of its associated electron density). Equation (2) defines the average scattering power ‘per atom’ (Guinier, 1994). If the total illuminated volume is larger than the coherence volume of the incident X-rays, and the coherence length of the scattering object is smaller than this coherence volume, then the observed scattering will be azimuthally isotropic, and the total coherent scattering intensity is given by the Debye scattering equation

$$\begin{aligned} I_{\text{coh}}(Q) &= \sum_{a_i, a_j \in \mathcal{A}} f_j^*(Q) f_i(Q) \frac{\sin Q r_{ij}}{Q r_{ij}} \\ &= \sum_{a_i \in \mathcal{A}} f_i^*(Q) f_i(Q) + \sum_{a_i \neq a_j \in \mathcal{A}} f_j^*(Q) f_i(Q) \frac{\sin Q r_{ij}}{Q r_{ij}} \quad (3) \\ &= N \langle f(Q)^2 \rangle + I_{\text{d}}(Q), \end{aligned}$$

where r_{ij} is the distance between atoms a_i and a_j . In the second equality, we have separated out the terms in the double summation for which $i = j$ —giving the self-scattering intensity, $N \langle f^2 \rangle$ —and for which $i \neq j$ —giving the so-called distinct scattering intensity I_{d} .

Only the latter term encodes the two-body information, so the reduced total scattering structure function, $F(Q)$, is simply the Q -weighted distinct scattering, normalized to the average scattering power of the sample, which is dual to the reduced PDF, $G(r)$, under the Fourier sine transform (Farrow & Billinge, 2009)

$$F(Q) := Q \frac{I_{\text{d}}(Q)}{N \langle f(Q)^2 \rangle} \overset{\mathcal{F}_s}{\leftrightarrow} G(r). \quad (4)$$

Although $F(Q)$ and $G(r)$ encode the same structural information, the PDF admits an intuitive interpretation, as it is essentially a weighted histogram of all the atomic pair distances in the sample.

2.2. Total scattering formalism for dilute solutions

In the differential solution scattering experiment, we perform two measurements. In measurement A [Fig. 1(a)], we collect the total scattering from a dilute solution (*e.g.* where the solute is a molecule); the second measurement B [Fig. 1(b)] is performed on an otherwise identical sample, lacking the solute (*i.e.* pure solvent plus any other constituents from A). Let V now denote the liquid unit cell of the solution, that

is, the minimal volumetric unit of composition of the homogeneous solution (Haldrup *et al.*, 2010), and \mathcal{A} denote the collection of atoms contained within this volume. We assume that the coherence volume, integrating over many liquid unit cells V , is identical in each sample (A and B), and that the scattering between distinct unit cells is incoherent. For example, in liquid water, the observed coherence length is $\lesssim 15 \text{ \AA}$ (Skinner *et al.*, 2013), which is much smaller than the typical coherence volume of synchrotron X-rays. Then it is sufficient to consider the scattering from just one such volume, averaged over all thermally accessible atomic configurations; hereafter, we will implicitly assume all quantities are ensemble-averaged, and thus azimuthally isotropic.

Considering sample A [Fig. 1(a)], we can subdivide the total collection of atoms $\{a_i\}_{i=1,\dots,N_{\mathcal{A}}} = \mathcal{A}$ into the set of solute atoms $u \in \mathcal{U}$ and the set of solvent atoms $v \in \mathcal{A} \setminus \mathcal{U}$. From equation (3), the total distinct scattering intensity from this volume element of measurement A can be decomposed

$$I_{\text{d}}^{\text{A}}(Q) = I_{uu}(Q) + I_{uv}(Q) + I_{vv}^{\text{A}}(Q), \quad (5)$$

where the first term is summed over solute–solute pairs, the second is summed over solute–solvent pairs and the final term is summed over the solvent–solvent pairs (Dohn *et al.*, 2015, 2023); the superscript in the final term will prove necessary to distinguish it from an analogous term arising from the set of all atoms contained in an equivalent volume element of sample B, denoted by \mathcal{B} . Indeed, for sample B, all of the atoms can be classified into solvent atoms $v \in \{a_i\}_{i=1,\dots,N_{\mathcal{B}}} = \mathcal{B}$, hence the distinct scattering from measurement B is simply

$$I_{\text{d}}^{\text{B}}(Q) = I_{vv}^{\text{B}}(Q). \quad (6)$$

The total ‘differential’ distinct scattering intensity arising from the atoms in V is thus

$$\Delta I_{\text{d}}(Q) := I_{\text{d}}^{\text{A}}(Q) - I_{\text{d}}^{\text{B}}(Q) = I_{uu}(Q) + I_{uv}(Q) + \Delta I_{\text{s}}(Q), \quad (7)$$

where $\Delta I_{\text{s}} := I_{vv}^{\text{A}} - I_{vv}^{\text{B}}$ arises from the contrast in the solvent–solvent correlations between the two measurements.

As we are interested in the solvation structure of the solute, it is natural to normalize ΔI_{d} to the scattering power of the solute $N_{\mathcal{U}} \langle f \rangle_{\mathcal{U}}^2$, in analogy to the standard formalism (the subscripts indicating that we consider only the set of solute atoms \mathcal{U}). This choice is also congruent with previously reported differential solution scattering experiments (Yang *et al.*, 2016; Xie *et al.*, 2022). We thus define the ‘differential’ reduced total scattering structure function by

$$\Delta F(Q) := Q \frac{\Delta I_{\text{d}}(Q)}{N_{\mathcal{U}} \langle f(Q) \rangle_{\mathcal{U}}^2} = F_{uu}(Q) + \tilde{F}_{uv}(Q) + \tilde{\Delta F}_{\text{s}}(Q), \quad (8)$$

which follows from equation (4) if we use $\tilde{\cdot}$ to denote that the corresponding term is normalized to the scattering power of the solute, rather than the scattering power of the atoms contributing to the interference function, as is usual [*c.f.* equation (4)].

The most immediate and obvious consequence of equation (8) is that $\Delta F \neq F_{uu}$, that is, in the differential setting, the derived interference function is not identical to the standard one as though we had measured the solute alone. Although

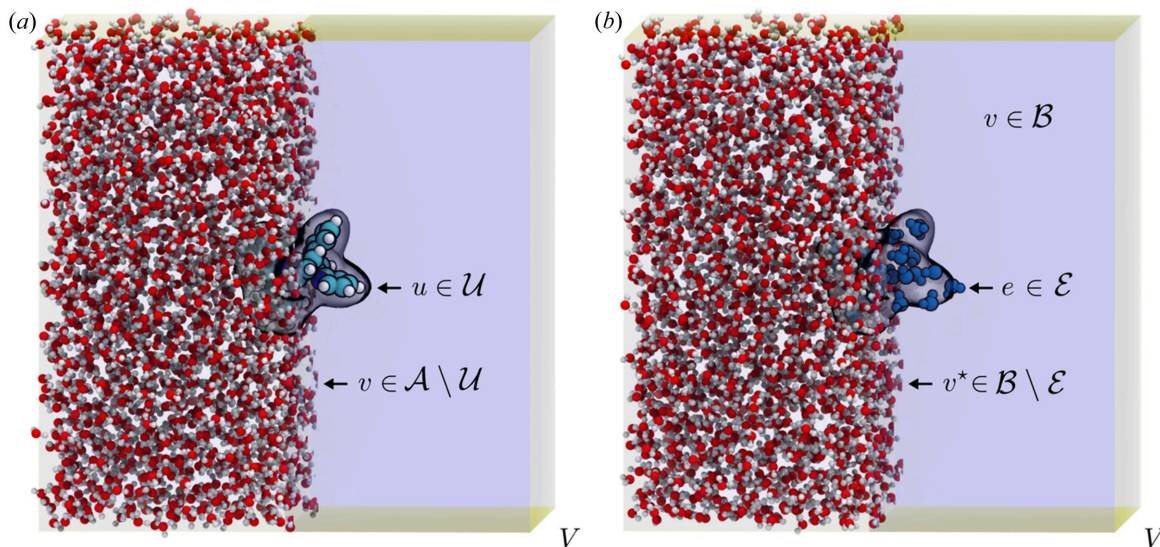


Figure 1 Description of the solution scattering experiment. (a) Schematic of the liquid unit cell V of sample A. The gray surface shows the van der Waals volume of the solute, defining the excluded volume. (b) Schematic of an equivalent volume V of sample B, showing how the excluded volume defines a droplet of solvent atoms (shown in blue). For clarity, only a half of the solvent atoms outside the excluded volumes are rendered.

F_{uu} is the dominant term in the differential (see below), the remaining two terms make non-negligible contributions to the total scattering differential. The solute–solvent term, \tilde{F}_{uv} , provides direct structural information pertaining to the local ordering of solvent about the solute. The differential solvent–solvent contribution, $\Delta\tilde{F}_s$, might be called the ‘solvent-restructuring’ term, in that it originates from the contrast between the solvent–solvent atomic pair-correlations of the two measurements. However, as we shall now show, this term is complicated by excluded-volume effects, which must be accounted for before interpreting $\Delta\tilde{F}_s$ in terms of solvent reorganization.

2.3. The solvent excluded volume

Referring to the schematics in Fig. 1, we recognize that, provided the solution behaves approximately ideally (*i.e.* the solvent–solvent interactions are unperturbed by the presence of the solute), the liquid unit cell of measurement A will always contain fewer solvent atoms than an equivalent volume of measurement B owing to the finite volume of the solute. Hence, the differential solvent–solvent structure function, $\Delta\tilde{F}_s$, must be corrected for this effect before interpreting how the solute affects the solvent structure. This is a necessary correction, because otherwise the solvent–solvent signals from the two samples are incommensurate, owing to the presence of excess solvent in sample B (or, equivalently, the absence of excluded solvent) (Soper, 1997).

Concretely, the surface of the van der Waals volume of the solute (V_U) demarcates the minimal volume of space where solvent is expected to be excluded in sample A, *i.e.* it defines the ‘excluded volume’ (Soper, 1997). Of course, in actuality, if the solute particle is interacting strongly with the solvent, it will affect the short-range interactions present between solvent molecules in the solution A, leading to a genuine

restructuring of the ‘bulk’ solvent. To assess the degree of this restructuring, $\Delta\tilde{F}_s$ should be corrected for the finite volume of the solute particle. Letting n_B denote the macroscopic atomic number density of the solvent in B, the number of excess (or excluded) solvent atoms in the volume V of sample B is given by $N_E = n_B V_U$; for each thermally accessible configuration of sample B, the excluded volume contains a ‘droplet’ of such atoms $e \in \mathcal{E}$, which, on average, sum to N_E atoms. This leads to a natural partitioning of the total set of solvent atoms in \mathcal{B} into \mathcal{E} and its complement $v^* \in \mathcal{B} \setminus \mathcal{E}$ [Fig. 1(b)].

In analogy to the decomposition of equation (5), we can express the distinct scattering from measurement B as

$$\begin{aligned} I_d^B(Q) &= I_{vv}^B(Q) = I_{ee}(Q) + I_{ev^*}(Q) + I_{v^*v^*}(Q) \\ &= I_x(Q) + I_{v^*v^*}(Q), \end{aligned} \quad (9)$$

where we have grouped the terms $I_{ee} + I_{ev^*} =: I_x$, given that both arise from correlations involving the **ex**cess or **ex**cluded solvent atoms. Note that total scattering experiments conducted in the HEXS regime typically do not collect data in the small-angle region (*i.e.* $Q_{\min} \gtrsim 1 \text{ \AA}^{-1}$) and are thus less sensitive to the nanoscopic morphology of the solvent excluded volume, as opposed to merely its size ($2\pi/Q_{\min} \simeq 10 \text{ \AA}$). However, *a priori*, it is not clear at which precise value of Q_{\min} the data become insensitive to the nanoscale details, hence, below we simply calculate I_x directly from atomistic models, as detailed in the Section S3.2 of the supporting information (see Section S3.2.1 of the supporting information for a more detailed discussion of the nanoscopic sensitivity of the excluded-volume term).

Following equations (7) and (8), we can thus expand the solvent–solvent differential

$$\begin{aligned}\Delta\tilde{F}_s(Q) &= \tilde{F}_{vv}^A(Q) - \tilde{F}_{vv}^B(Q) \\ &= \left[\tilde{F}_{vv}^A(Q) - \tilde{F}_{v^*,v^*}(Q) \right] - \tilde{F}_x(Q) \quad (10) \\ &= \Delta\tilde{F}_r(Q) - \tilde{F}_x(Q).\end{aligned}$$

The physical meaning of the remaining differential, $\Delta\tilde{F}_r := \tilde{F}_{vv}^A - \tilde{F}_{v^*,v^*}$, can now be properly understood by ‘solvent restructuring’ in that it describes how the introduction of the solute changes the solvent structure ‘outside of the excluded-volume region’.

Another way to appreciate the consistency of this concept is to consider the asymptotic limits of each term in equation (10) as a function of the volume fraction of the solute, v . As $v \rightarrow 0$, the excluded volume vanishes, and both $\tilde{F}_x \rightarrow \mathbf{0}$ and $\Delta\tilde{F}_r \rightarrow \mathbf{0}$, implying $\Delta\tilde{F}_s = \mathbf{0}$; sample A becomes identical to sample B. As $v \rightarrow 1$, there is no longer any solvent remaining in sample A, hence, tautologically, there can be no solvent reorganization ($\Delta\tilde{F}_r \rightarrow \mathbf{0}$); however, $\tilde{F}_{vv}^B \rightarrow \tilde{F}_x \neq \mathbf{0}$, and thus $\Delta\tilde{F}_s \rightarrow -\tilde{F}_x \neq \mathbf{0}$. In this limit, all of the solvent becomes ‘excluded’.

Although these are pure *gedanken* experiments, they serve to illustrate the necessity of separating out the excluded solvent structure function from the total differential solvent–solvent term before assessing the importance of solvent reorganization in a real experiment. At any finite v , both \tilde{F}_x and $\Delta\tilde{F}_r$ can be expected to contribute to the total differential signal, although for solutes that interact weakly with the solvent (*i.e.* ideal–dilute solutions), $\Delta\tilde{F}_r \approx \mathbf{0}$ (see below).

2.4. Summary of the framework

To summarize, we are concerned with the structural information encoded in a differential total scattering experiment, $\Delta I(Q) := I^A(Q) - I^B(Q)$. In Section 2.2, we arrived at an exact formulation for the differential reduced total scattering structure function, $\Delta F(Q)$, equation (8). Though this is an exact result, we have further introduced the concept of the solvent excluded volume in order to decompose the differential solvent–solvent structure function, $\Delta\tilde{F}_s$, into the sum of two contributions, one arising from the fact that the solute displaces a finite volume of solvent upon dissolution (\tilde{F}_x), and one arising from genuine restructuring of the solvent upon perturbation by the solute ($\Delta\tilde{F}_r$).

Thus, combining equations (8) and (10), we arrive at a parameterization of the differential reduced structure function

$$\Delta F(Q) = F_{uu}(Q) + \tilde{F}_{uv}(Q) + \Delta\tilde{F}_r(Q) - \tilde{F}_x(Q). \quad (11)$$

Since the notion of the solvent excluded volume is imposed *a priori* to decompose the pair-correlations from the pure solvent measurement B, equation (11) represents our proposed model for representing the information in the differential structure function. While we have motivated this decomposition physically, by reference to the finite volume of the solute in measurement A, the solvent excluded volume remains an imprecise, if useful, definition, depending as it does on a choice for computing the solute van der Waals volume.

The first three terms in equation (11) correspond to the solvation structure of the solute, understood with respect to the intra- and intermolecular interactions present in the solution, and its deviation from ideal-dilute behavior. The solute–solute structure function (F_{uu}) encodes the intramolecular structure of the solute in solution, averaged over the thermal ensemble. The solute–solvent structure function (\tilde{F}_{uv}) encodes the local ordering of solvent atoms about the solute, recovering the usual notion of the ‘solvation cage’ (Kim *et al.*, 2009; Ihee, 2009; Haldrup *et al.*, 2010). These terms are measured directly from the solution sample A. The solvent-restructuring structure function ($\Delta\tilde{F}_r$) is defined by the contrast between measurements A and B, and in this sense can be considered an emergent function of the differential experiment. While $\Delta\tilde{F}_r$ does not directly encode any two-body information regarding the solute atoms, it provides a concrete measure of the ideality of the solution, from a structural perspective (Martin & Shipman, 2023). By definition, in an ideal-dilute solution, the solvent–solvent interactions are unaffected by the presence of the solute and, hence, can be identified with the case that $\Delta\tilde{F}_r = \mathbf{0}$.

The excluded solvent structure function (\tilde{F}_x), however, represents an essentially artifactual term arising from the differential approach, in that it does not encode any two-body information from the solution sample A. Although it is defined in reference to the solute structure, as a function of V_U , the inversion of this mapping (*i.e.* obtaining V_U from \tilde{F}_x) is likely to be better addressed via SAXS (*c.f.* Section 3.2.1 of the supporting information) (Jeffries *et al.*, 2021). As such, \tilde{F}_x can be considered a ‘necessary evil’ of the differential solution HEXS experiment.

Finally, the ‘differential reduced PDF’ is naturally defined to be the Fourier sine transform of $\Delta F(Q)$ which, owing to the linearity of the transform, gives

$$\begin{aligned}\Delta G(r) &:= \mathcal{F}_s[\Delta F(Q)](r) \\ &= G_{uu}(r) + \tilde{G}_{uv}(r) + \Delta\tilde{G}_r(r) - \tilde{G}_x(r), \quad (12)\end{aligned}$$

where, again, we use $\tilde{}$ to remind us that the corresponding functions are transformed from renormalized structure functions. Equation (12) emphasizes the importance of a detailed theoretical understanding of the information content of the differential solution scattering experiment. If, in deriving a PDF from such an experiment, one assumes that subtracting a pure solvent reference from the raw solution scattering signal eliminates all solvent–solvent correlations, one might erroneously attribute features arising from $\Delta\tilde{G}_r$ or \tilde{G}_x to the solute structure. Modeling an experimentally obtained PDF with G_{uu} alone (Megyes *et al.*, 2003, 2005; Neufeind *et al.*, 2004), or indeed using a ‘dummy atom’ approach to account for the excluded volume (Svergun *et al.*, 1995; Tiede *et al.*, 2004; Tiede *et al.*, 2009), may serve for an approximate description of the intramolecular structure of the solute. To achieve a quantitative description of the solvation structure, a more rigorous model is required (see below).

3. Numerical evaluation

To evaluate our model numerically, we simulated the differential solution scattering experiment using molecular dynamics (MD). By explicitly constructing two ensembles corresponding to samples A and B, we can perform an exact MD experiment [within the accuracy of the force field parameterization and the numerical precision in evaluating equation (8)] and analyze the results in terms of equation (11). As a test system, we considered aqueous solutions of tris(2,2'-bipyridine)ruthenium(II) dichloride ($[\text{Ru}(\text{bpy})_3]\text{Cl}_2$), a classic transition metal complex whose solution structure has been the subject of numerous theoretical investigations (Moret *et al.*, 2009; Josefsson *et al.*, 2016; Abedi *et al.*, 2019). See the supporting information for complete details regarding the MD simulations and subsequent calculation of the coherent scattering intensities.

3.1. Construction

We first simulated sample A as a 15 mM aqueous solution of $[\text{Ru}(\text{bpy})_3]\text{Cl}_2$ by solvating a single solute molecule in a box containing 3535 water molecules, and performing MD in the isothermal-isobaric (NpT) ensemble for 100 ns at 300 K and 1 atm (in five independent 20 ns trajectories). To properly simulate sample B, we must determine the (ensemble-averaged) number of excluded solvent atoms, N_E . A simple heuristic procedure is to estimate $N_E \approx n_B V_U$ (see above), where n_B is the (experimental) atomic number density of the solvent (for water, $n_B = 0.1$ atoms \AA^{-3}). Using this heuristic, we find $N_E = 66$ atoms, or 22 water molecules, hence sample B was simulated by performing MD on a box of 3557 water molecules for 20 ns in the NpT ensemble at 300 K and 1 atm.

To investigate the limits of the dilute solution phase for this system, we performed an additional solution simulation at 0.5 M, constructed from the pure solvent reference by replacing 22 water molecules for each copy of $[\text{Ru}(\text{bpy})_3]\text{Cl}_2$ included in the simulation cell (32). We found that this system behaves identically up to 0.5 M (see the supporting information for further discussion), and restrict our discussion to the 15 mM simulation hereafter as this condition matches our experimental results below.

3.2. Evaluating the model

As expected, the total coherent scattering intensity computed from the liquid unit cell of a 15 mM solution of $[\text{Ru}(\text{bpy})_3]\text{Cl}_2$ is nearly indistinguishable, by eye, from that computed for an equivalent volume of neat water [Fig. 2(a), top]. The differential scattering intensity exhibits both a smoothly decreasing background, arising from the differential self-scattering, as well as the oscillatory features arising from the differential interference function (ΔI_d), and is about 100 times weaker than the total coherent scattering intensities [Fig. 2(a), bottom]. Normalizing the interference function to the scattering power of $[\text{Ru}(\text{bpy})_3]\text{Cl}_2$ and weighting the result by Q produces the differential reduced structure function ΔF defined in equation (8) [Fig. 2(b)].

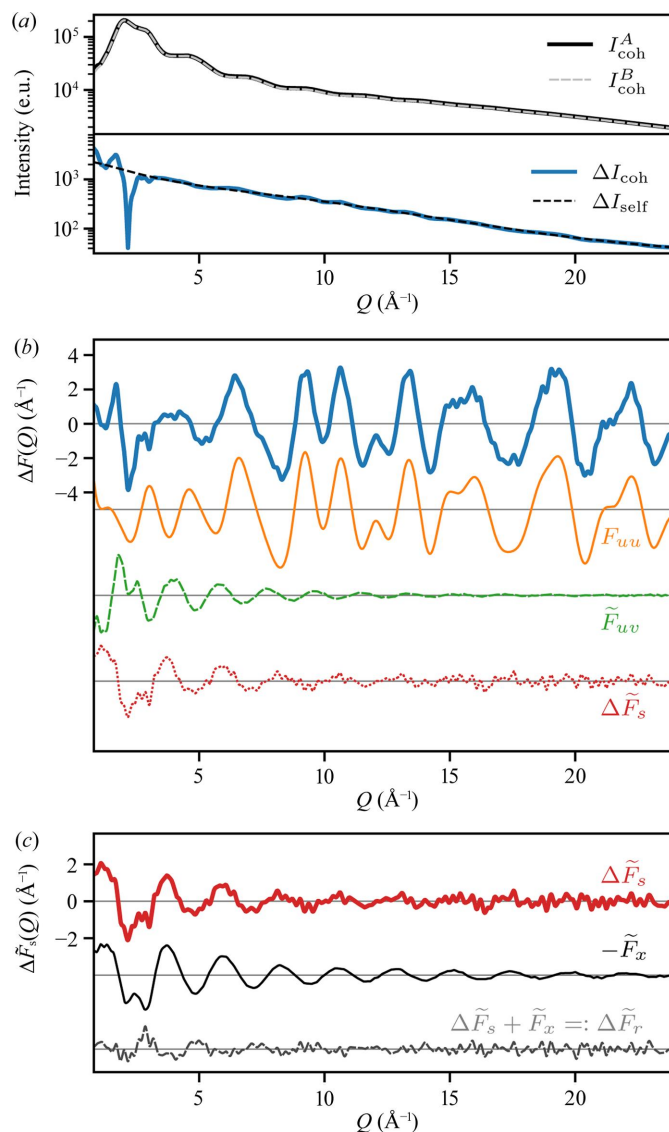


Figure 2

Simulation of the differential solution scattering experiment. (a) Top, total coherent scattering intensity from an MD simulation of a 15 mM aqueous solution of $[\text{Ru}(\text{bpy})_3]\text{Cl}_2$ (I_{coh}^A , solid black line) and a corresponding simulation of neat water (I_{coh}^B , dashed gray line); bottom, the total differential coherent scattering intensity (ΔI_{coh} , solid blue line) along with the differential self-scattering (ΔI_{self} , dashed black line). (b) The differential reduced structure function, ΔF (solid blue line), and its decomposition into F_{uu} (solid orange line), \tilde{F}_{uv} (dashed green line) and $\Delta \tilde{F}_s$ (dotted red line). (c) Further decomposition of $\Delta \tilde{F}_s$ (solid red line) into $-\tilde{F}_x$ (solid black line) and $\Delta \tilde{F}_r$ (dashed gray line).

Fig. 2(b) also shows the contributions to ΔF from each term in equation (8), illustrating that, while F_{uu} is in general large throughout reciprocal space (as a result of the rigid intramolecular structure of $[\text{Ru}(\text{bpy})_3]^{2+}$), the remaining terms make non-negligible contributions at low- to moderate- Q . For example, the solute-solvent structure function (\tilde{F}_{uv}) produces a strong peak at $Q \approx 1.7 \text{ \AA}^{-1}$, reminiscent of the ‘first sharp diffraction peak’ often observed in network-forming liquids and glasses (Elliott, 1995). The differential solvent-solvent structure function ($\Delta \tilde{F}_s$) noticeably modulates the intensity throughout $Q \in [1, 7] \text{ \AA}^{-1}$.

To further decompose $\Delta\tilde{F}_s$ into the excluded solvent and solvent-restructuring terms, we performed an independent calculation of \tilde{F}_x , and take $\Delta\tilde{F}_r = \Delta\tilde{F}_s + \tilde{F}_x$, by definition, following equation (10). To compute \tilde{F}_x , we adopt the following strategy. For each particle of the solute (*i.e.* $[\text{Ru}(\text{bpy})_3]^{2+}$ and two solvated Cl^- counterions), we use a representative molecular structure (*e.g.* for $[\text{Ru}(\text{bpy})_3]^{2+}$, we used the ground-state structure optimized via DFT), and for each frame of simulation B we embed this structure into the center of the simulation. For a given frame, we find all solvent atoms that lie within one van der Waals radius of any atom of the solute particle, and label them as belonging to the excluded solvent droplet. In this fashion, all solvent atoms in all frames of the trajectory are labeled as belonging to the set $e \in \mathcal{E}$ or the set $v^* \in \mathcal{B} \setminus \mathcal{E}$ [*c.f.* Fig. 1(b)]. The total coherent scattering intensity calculated from simulation B is thus decomposed in accordance with equation (9). Note, therefore, that our evaluation of the solvent-restructuring term, $\Delta\tilde{F}_r$, is dependent on the precise algorithm used to perform this decomposition (*c.f.* Section S3.2 of the supporting information). However, the approach described above produces self-consistent results with our semi-empirical heuristic for $N_{\mathcal{E}}$ used in construction of the simulation, finding that $N_{\mathcal{E}} = 66$ atoms in the ensemble-average.

As shown in Fig. 2(c), $\Delta\tilde{F}_s$ is dominated by the excluded solvent structure function (\tilde{F}_x). The amplitude of the solvent-restructuring structure function ($\Delta\tilde{F}_r$) damps nearly to zero by $Q \approx 6 \text{ \AA}^{-1}$, while the ratio $\|\tilde{F}_x\|_{\ell_2}/\|\Delta\tilde{F}_r\|_{\ell_2} = 2.4$ over $Q \in [0.6, 24] \text{ \AA}^{-1}$. This numerical result indicates that, at least within the accuracy of our simulation, and our algorithm for computing \tilde{F}_x , $[\text{Ru}(\text{bpy})_3]\text{Cl}_2$ does not significantly restructure water. Without a more detailed investigation, beyond our scope here, we also caution that the sensitivity of $\Delta\tilde{F}_s$ to restructuring (*i.e.* the region where $\Delta\tilde{F}_r$ is different from zero) is likely to be localized to the low- Q region, which is typically truncated in data collection optimized for HEXS-based PDF studies (*c.f.* Section S3.2.1 of the supporting information).

Now we wish to understand the relative importance of each term summing to our complete model. For this, we require a statistical measure of the contribution of each of these terms to approximate the total result of equation (11). As our focus is on developing a PDF formalism, we will perform this analysis on the real space analogs [equation (12)]. Considering ΔG as ground truth, one measure of the explanatory power of a model is given by the coefficient of determination (r^2) which ranges from 0 (baseline model) to 1 (perfect linear model); alternatively, we can judge the goodness-of-fit of each model via the standard crystallographic R factor. Table 1 lists both metrics for a series of models given by sequential addition of each term in equation (12); these same data are presented graphically in Fig. 3.

From these results, it is clear that G_{uu} dominates the total PDF, as expected ($r^2 = 0.84$; $R = 40\%$). However, to achieve near quantitative accuracy, it is necessary to consider both \tilde{G}_{uv} as well as \tilde{G}_x ($r^2 = 0.99$; $R = 12\%$). Indeed, in terms of accuracy, the excluded solvent structure matters more than the solute-solvent structure, despite its ‘artificial’ nature. In this system,

Table 1
Statistical decomposition of the framework.

Model	r^2	R factor (%)
G_{uu}	0.84	40
$G_{uu} + \tilde{G}_{uv}$	0.89	33
$G_{uu} + \tilde{G}_{uv} - \tilde{G}_x$	0.99	12
$G_{uu} + \tilde{G}_{uv} - \tilde{G}_x + \Delta\tilde{G}_r = \Delta G$	1.00	0.0

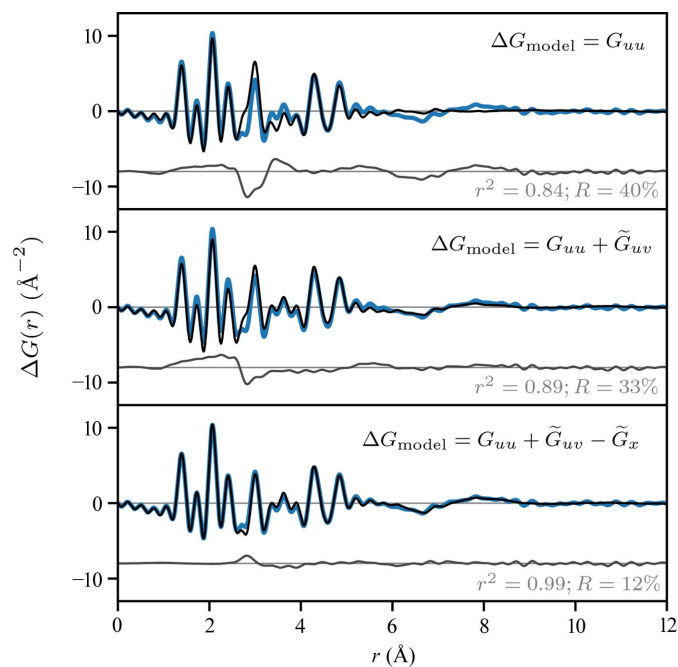


Figure 3
Various approximations to ΔG . Each plot compares the exact ΔG (solid blue line) from the simulated experiment with various models (ΔG_{model} , solid black line), as indicated, of increasing complexity. The residuals are shown below in gray. Data are the Fourier transforms of those in Fig. 2(b) using $Q \in [0.8, 21.4] \text{ \AA}^{-1}$.

the solvent restructuring appears nearly negligible (see above); the R factor for a model neglecting this term (12%) is already about a factor of 2 to 3 times lower than typically achieved in modeling PDFs of molecular systems (20–30%) (Terban & Billinge, 2022). Nevertheless, in general the importance of solvent restructuring should be assessed on a case-by-case basis, by independent thermodynamic measures of solution ideality, for example.

4. Experimental validation

To validate our framework against experimental data, we performed total scattering measurements at beamline 11-ID-B of the Advanced Photon Source. We collected data on a 15 mM aqueous solution of $[\text{Ru}(\text{bpy})_3]\text{Cl}_2$ and an equivalent sample of neat water in the range $Q \in [0.6, 24] \text{ \AA}^{-1}$, and reduced these data to ΔF . In practice, this requires access to flux-normalized total scattering data, which are often not available with the required precision at PDF-focused beamlines such as 11-ID-B. Typically, the solution to this problem would be to obtain absolutely normalized total scattering data during reduction, using the high- Q behavior of the raw data to

compute normalization terms, as it is dominated by the structure-independent Compton scattering.

However, since this requires high-quality data collected out to high- Q and accurate calculation of the Compton backgrounds (which in turn demands rigorous precision in the preparation of dilute solutions), we detail an alternative approach for data reduction that automates this task in the supporting information. Briefly, from the derivation of the theoretical framework above, it is straightforward to perform data normalization and subtraction simultaneously, utilizing the standard information required for reduction of total scattering data to a PDF, but incorporating priors regarding the solvent excluded volume and the ideality of the solution to render the reduction a convex program. See Section S4 of the supporting information for full details on the data reduction algorithm.

The experimental ΔF is shown in Fig. 4(a), alongside our MD calculation. Qualitatively, the agreement is excellent: every feature observed in the experiment is reproduced by the calculation. The residual is dominated by modulations that persist throughout reciprocal space, indicating that the largest source of error in our simulation lies in the intramolecular solute–solute structure. This is borne out in real-space analysis [Fig. 4(b)], where the majority of the residual arises from small errors in the peak profiles below $r \simeq 5$ Å, corresponding to the sharp, intramolecular features attributable to $[\text{Ru}(\text{bpy})_3]^{2+}$. Significantly, at $r \gtrsim 5$ Å, the residual drops to the level of the Fourier truncation ripples in the experimental data. As the features at long- r arise from the subtle combination of the solute–solvent, excluded solvent and solvent-restructuring terms (Fig. 3), the quantitative agreement in this region of real space indicates that our simulation accurately captures these structural features.

Given the difficulty in constructing accurate molecular mechanics force fields for coordination complexes (Deeth, 2010), the errors in our simulated F_{uu}/G_{uu} are unsurprising. Nevertheless, the overall level of agreement is remarkable considering that no attempt has been made here to optimize the simulation against the experimental data. Indeed, the most significant conclusion to be drawn from these results is that our framework provides a realistic and accurate parameterization of the differential solution scattering experiment. Such a parameterization is a necessary prerequisite for solving the inverse modeling problem (*i.e.* solving the solution structure from experimental scattering data).

To solve this inverse problem, one could, for example, use the experimental PDF to drive the optimization of the force field (Hsu *et al.*, 2020). Alternatively, if classical force fields are unsuitable (a likely scenario for transition metal complexes), one could use quantum-mechanically derived normal modes to parameterize the internal degrees of freedom of the solute (*i.e.* F_{uu}), and combine this with a coarser-grained MD parameterization for the remaining terms (Leshchev *et al.*, 2023). Discussion of the structural features in the experimental scattering data of $[\text{Ru}(\text{bpy})_3]\text{Cl}_2$, in terms of a refined theoretical model, is beyond the scope of this work, and will be presented in a subsequent publication. Here, we simply note

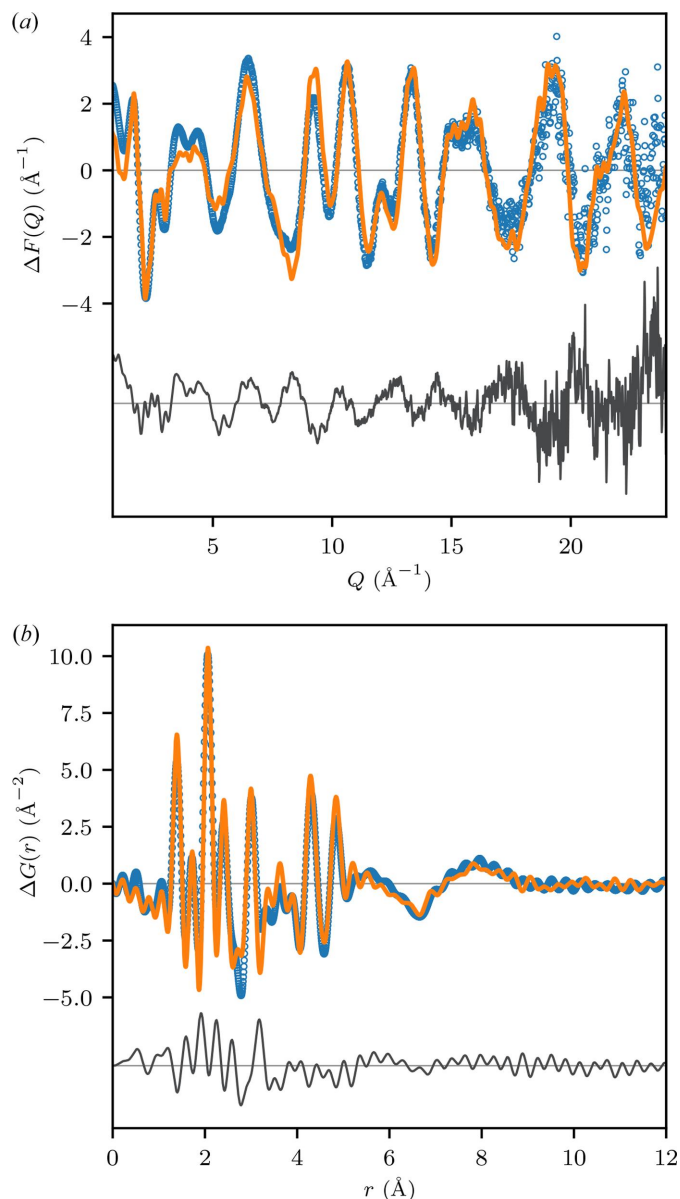


Figure 4
Comparison of the experimental solution scattering data with MD simulation. (a) Comparison of the experimental ΔF (open blue circles) derived from a 15 mM aqueous solution of $[\text{Ru}(\text{bpy})_3]\text{Cl}_2$ with calculation (solid orange line); the residual is shown below in gray. (b) Data from (a) Fourier transformed to real space using $Q \in [0.8, 21.4] \text{ \AA}^{-1}$.

that both our simulated and experimental results are in broad agreement with prior theoretical studies of the aqueous solvation structure of $[\text{Ru}(\text{bpy})_3]\text{Cl}_2$, in particular in the presence of two resolved solvation shells with peak maxima near 5.5 and 8.0 Å [Fig. 4(b)] (Moret *et al.*, 2009; Josefsson *et al.*, 2016; Abedi *et al.*, 2019).

5. Conclusions

Although solvation can be a critical determinant of chemistry in the condensed phases, few experimental techniques are capable of determining the structure of molecules in solution with sub-atomic resolution. In principle, total X-ray scattering

measurements yield such information, but are complicated by the necessity of a differential analysis to extract the relevant scattering data against a large experimental background. In order to fully understand the information content of such experiments, we have developed a theoretical framework for differential solution scattering, provided by equations (11) and (12). This framework was evaluated with the aid of numerical simulations and validated by comparison to experiment.

Our results lay the groundwork for developing quantitative simulations of solvation structure, as measured by total X-ray scattering. Efforts directed at evaluating quantitative simulation methodologies are ongoing at Argonne. Ultimately, one can envision that solution scattering measurements will provide insights into atomic structure with accuracy and precision approaching crystallography, but instead of reporting on the average (macroscopic) structure of a molecule in a crystal, elucidating the local (nanoscopic) structure of a molecule in complex condensed phases.

While our focus has been on solvation structure, the application of total scattering measurements to complex, multiphase systems is increasingly popular (Jensen *et al.*, 2015; Roelsgaard *et al.*, 2019; He *et al.*, 2020; Hoffman *et al.*, 2024). In these settings, challenges similar to those of the solution scattering experiment arise (*e.g.* a large experimental background from a bulk substrate onto which a thin film of interest is deposited). Though not immediately applicable, we believe that the methodology described here should provide a useful framework for understanding the data acquired in any total scattering experiment where a differential analysis proves necessary.

6. Related literature

The following references are cited in the supporting information: Frisch *et al.* (2016); Staroverov *et al.* (2003); Weigend & Ahlrichs (2005); Singh & Kollman (1984); Besler *et al.* (1990); Rappe *et al.* (1992); Li & Merz (2016); Case *et al.* (2021); Izadi *et al.* (2014); Sengupta *et al.* (2021); Eastman *et al.* (2017); Warren (1990); Neder & Proffen (2020); Juhás *et al.* (2015); Waasmaier & Kirfel (1995); McGibbon *et al.* (2015); Thijsse (1984); Cromer & Mann (1967); Toby & Von Dreele (2013); Peterson *et al.* (2021); Qiu *et al.* (2004); Weng *et al.* (2023); Juhás *et al.* (2013).

Acknowledgements

The authors thank J. M. Hoffman, O. J. Borkiewicz, G. Kwon and T. W. Kim for helpful discussions.

Funding information

We acknowledge support for this work from the US Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, Division of Chemical, Biological and Geological Science through Argonne National Laboratory (ANL) (contract No. DE-AC02-06CH11357). High-energy

X-ray scattering measurements were carried out at beamline 11-ID-B of the Advanced Photon Source, an Office of Science User Facility operated for the US DOE, Office of Science, by ANL and supported by the US DOE (contract No. DE-AC02-06CH11357).

References

- Abedi, M., Levi, G., Zederkof, D. B., Henriksen, N. E., Pápai, M. & Møller, K. B. (2019). *Phys. Chem. Chem. Phys.* **21**, 4082–4095.
- Attanasio, D. (1986). *J. Phys. Chem.* **90**, 4952–4957.
- Besler, B. H., Merz, K. M. Jr & Kollman, P. A. (1990). *J. Comput. Chem.* **11**, 431–439.
- Biasin, E., Fox, Z. W., Andersen, A., Ledbetter, K., Kjaer, K. S., Alonso-Mori, R., Carlstad, J. M., Chollet, M., Gaynor, J. D., Glowia, J. M., Hong, K., Kroll, T., Lee, J. H., Liekhus-Schmaltz, C., Reinhard, M., Sokaras, D., Zhang, Y., Doumy, G., March, A. M., Southworth, S. H., Mukamel, S., Gaffney, K. J., Schoenlein, R. W., Govind, N., Cordones, A. A. & Khalil, M. (2021). *Nat. Chem.* **13**, 343–349.
- Bifulco, G., Dambruoso, P., Gomez-Paloma, L. & Riccio, R. (2007). *Chem. Rev.* **107**, 3744–3779.
- Billinge, S. J. L. (2019). *Philos. Trans. R. Soc. A.* **377**, 20180413.
- Case, D. A., Aktulga, H. M., Belfon, K., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, T. E. III, Cisneros, G. A., Cruzeiro, V. W. D., Darden, T. A., Duke, R. E., Giambasu, G., Gilson, M. K., Gohlke, H., Goetz, A. W., Harris, R., Izadi, S., Izmailov, S. A., Jin, C., Kasavajhala, K., Kaymak, M. C., King, E., Kovalenko, A., Kurtzman, T., Lee, T. S., LeGrand, S., Li, P., Lin, C., Liu, J., Luchko, T., Luo, R., Machado, M., Man, V., Manathunga, M., Merz, K. M., Miao, Y., Mikhailovskii, O., Monard, G., Nguyen, H., O’Hearn, K. A., Onufriev, A., Pan, F., Pantano, S., Qi, R., Rahnamoun, A., Roe, D. R., Roitberg, A., Sagui, C., Schott-Verdugo, S., Shen, J., Simmerling, C. L., Skrynnikov, N. R., Smith, J., Swails, J., Walker, R. C., Wang, J., Wei, H., Wolf, R. M., Wu, X., Xue, Y., York, D. M., Zhao, S. & Kollman, P. A. (2021). *Amber2021*, University of California, San Francisco.
- Chen, L. X. (2004). *Angew. Chem. Int. Ed.* **43**, 2886–2905.
- Chen, P. & Hub, J. S. (2014). *Biophys. J.* **107**, 435–447.
- Cheng, H., Sun, Q., Li, L., Zou, Y., Wang, Y., Cai, T., Zhao, F., Liu, G., Ma, Z., Wahyudi, W., Li, Q. & Ming, J. (2022). *ACS Energy Lett.* **7**, 490–513.
- Clore, G. M. & Gronenborn, A. M. (1991). *Science*, **252**, 1390–1399.
- Cromer, D. T. & Mann, J. B. C. (1967). *J. Chem. Phys.* **47**, 1892–1893.
- Deeth, R. J. (2010). *Molecular Mechanics for Transition Metal Centers: From Coordination Complexes To Metalloproteins*, vol. 62, pp. 1–39. Academic Press. <https://www.sciencedirect.com/science/article/pii/S0898883810620016>.
- Dohn, A. O., Biasin, E., Haldrup, K., Nielsen, M. M., Henriksen, N. E. & Møller, K. B. (2015). *J. Phys. B At. Mol. Opt. Phys.* **48**, 244010.
- Dohn, A. O., Markmann, V., Nimmrich, A., Haldrup, K., Møller, K. B. & Nielsen, M. M. (2023). *J. Chem. Phys.* **159**, 124115.
- Driel, T. B. van, Kjaer, K. S., Hartsock, R. W., Dohn, A. O., Harlang, T., Chollet, M., Christensen, M., Gawelda, W., Henriksen, N. E., Kim, J. G., Haldrup, K., Kim, K. H., Ihee, H., Kim, J., Lemke, H., Sun, Z., Sundström, V., Zhang, W., Zhu, D., Møller, K. B., Nielsen, M. M. & Gaffney, K. J. (2016). *Nat. Commun.* **7**, 13678.
- Drover, M. W. (2022). *Chem. Soc. Rev.* **51**, 1861–1880.
- Eastman, P., Swails, J., Chodera, J. D., McGibbon, R. T., Zhao, Y., Beauchamp, K. A., Wang, L.-P., Simmonett, A. C., Harrigan, M. P., Stern, C. D., Wiewiora, R. P., Brooks, B. R. & Pande, V. S. (2017). *PLoS Comput. Biol.* **13**, e1005659.
- Egami, T. & Billinge, S. J. (2012). *Underneath the Bragg Peaks: Structural Analysis of Complex Materials*. 2nd ed. New York: Pergamon.
- Elliott, S. (1995). *J. Non-Cryst. Solids*, **182**, 40–48.

- Farrow, C. L. & Billinge, S. J. L. (2009). *Acta Cryst.* **A65**, 232–239.
- Filippini, A. (1994). *J. Phys. Condens. Matter*, **6**, 8415–8427.
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Petersson, G. A., Nakatsuji, H., Li, X., Caricato, M., Marenich, A. V., Bloino, J., Janesko, B. G., Gomperts, R., Mennucci, B., Hratchian, H. P., Ortiz, J. V., Izmaylov, A. F., Sonnenberg, J. L., Williams-Young, D., Ding, F., Lipparini, F., Egidi, F., Goings, J., Peng, B., Petrone, A., Henderson, T., Ranasinghe, D., Zakrzewski, V. G., Gao, J., Rega, N., Zheng, G., Liang, W., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Throssell, K., Montgomery, J. A. Jr, Peralta, J. E., Ogliaro, F., Bearpark, M. J., Heyd, J. J., Brothers, E. N., Kudin, K. N., Staroverov, V. N., Keith, T. A., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A. P., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Millam, J. M., Klene, M., Adamo, C., Cammi, R., Ochterski, J. W., Martin, R. L., Morokuma, K., Farkas, O., Foresman, J. B. & Fox, D. J. (2016). *Gaussian 16*. Revision C. 01, Gaussian Inc. Wallingford CT.
- Garino, C., Borfecchia, E., Gobetto, R., van Bokhoven, J. A. & Lamberti, C. (2014). *Coord. Chem. Rev.* **277–278**, 130–186.
- Ghosh, A., Dasgupta, S., Kundu, A. & Mandal, S. (2022). *Dalton Trans.* **51**, 10320–10337.
- Grübel, G. & Zontone, F. (2004). *J. Alloys Compd.* **362**, 3–11.
- Guinier, A. (1994). *X-ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies*. New York: Dover Publications.
- Haldrup, K., Christensen, M. & Meedom Nielsen, M. (2010). *Acta Cryst.* **A66**, 261–269.
- He, X., Waldman, R. Z., Mandia, D. J., Jeon, N., Zaluzec, N. J., Borkiewicz, O. J., Ruett, U., Darling, S. B., Martinson, A. B. F. & Tiede, D. M. (2020). *ACS Nano*, **14**, 14846–14860.
- Hertz, H. G. (1970). *Angew. Chem. Int. Ed. Engl.* **9**, 124–138.
- Hoffman, J. M., Thompson, N. B., Borkiewicz, O., He, X., Amsterdam, S., Xie, Z., -, Taggart, A., Mulfort, K. L., Martinson, A. B. F., Chen, L. X., Ruett, U. & Tiede, D. M. (2024). *IUCrJ*, **11**, 120–128.
- Hsu, D. J., Leshchev, D., Kosheleva, I., Kohlstedt, K. L. & Chen, L. X. (2020). *J. Chem. Phys.* **152**, 204115.
- Ihee, H. (2009). *Acc. Chem. Res.* **42**, 356–366.
- Izadi, S., Anandakrishnan, R. & Onufriev, A. V. (2014). *J. Phys. Chem. Lett.* **5**, 3863–3871.
- Jay, R. M., Vaz da Cruz, V., Eckert, S., Fondell, M., Mitzner, R. & Föhlisch, A. (2020). *J. Phys. Chem. B*, **124**, 5636–5645.
- Jeffries, C. M., Ilavsky, J., Martel, A., Hinrichs, S., Meyer, A., Pedersen, J. S., Sokolova, A. V. & Svergun, D. I. (2021). *Nat. Rev. Methods Primers*, **1**, 70.
- Jensen, K. M. Ø., Blichfeld, A. B., Bauers, S. R., Wood, S. R., Door- yhée, E., Johnson, D. C., Iversen, B. B. & Billinge, S. J. L. (2015). *IUCrJ*, **2**, 481–489.
- Josefsson, I., Eriksson, S. K., Rensmo, H. & Odelius, M. (2016). *Struct. Dyn.* **3**, 023607.
- Juhás, P., Davis, T., Farrow, C. L. & Billinge, S. J. L. (2013). *J. Appl. Cryst.* **46**, 560–566.
- Juhás, P., Farrow, C., Yang, X., Knox, K. & Billinge, S. (2015). *Acta Cryst.* **A71**, 562–568.
- Katayama, T., Choi, T.-K., Khakhulin, D., Dohn, A. O., Milne, C. J., Vankó, G., Németh, Z., Lima, F. A., Szlachetko, J., Sato, T., Nozawa, S., Adachi, S., Yabashi, M., Penfold, T. J., Gawelda, W. & Levi, G. (2023). *Chem. Sci.* **14**, 2572–2584.
- Khakhulin, D., Lawson Daku, L. M., Leshchev, D., Newby, G. E., Jarenmark, M., Bressler, C., Wulff, M. & Canton, S. E. (2019). *Phys. Chem. Chem. Phys.* **21**, 9277–9284.
- Kim, T. K., Lee, J. H., Wulff, M., Kong, Q. & Ihee, H. (2009). *ChemPhysChem*, **10**, 1958–1980.
- Köfinger, J. & Hummer, G. (2013). *Phys. Rev. E*, **87**, 052712.
- Leshchev, D. J. S., Valentine, A., Kim, P., Mills, A. W., Roy, S., Chakraborty, A., Biasin, E., Haldrup, K., Hsu, D. J., Kirschner, M. S., Rimmerman, D., Chollet, M., Glowonia, J. M., van Driel, T. B., Castellano, F. N., Li, X. & Chen, L. X. (2023). *Angew. Chem. Int. Ed.* **62**, e202304615.
- Li, P. & Merz, K. M. J. (2016). *J. Chem. Inf. Model.* **56**, 599–604.
- Marcus, R. A. (1993). *Angew. Chem. Int. Ed. Engl.* **32**, 1111–1121.
- Martin, J. D. & Shipman, A. M. (2023). *J. Electrochem. Soc.* **170**, 066508.
- McGibbon, R. T., Beauchamp, K. A., Harrigan, M. P., Klein, C., Swails, J. M., Hernández, C. X., Schwantes, C. R., Wang, L.-P., Lane, T. J. & Pande, V. S. (2015). *Biophys. J.* **109**, 1528–1532.
- Megyes, T., Bálint, S., Bakó, I., Grósz, T. & Pálinkás, G. (2008). *J. Am. Chem. Soc.* **130**, 9206–9207.
- Megyes, T., Jude, H., Grósz, T., Bakó, I., Radnai, T., Tárkányi, G., Pálinkás, G. & Stang, P. J. (2005). *J. Am. Chem. Soc.* **127**, 10731–10738.
- Megyes, T., Schubert, G., Kovács, M., Radnai, T., Grósz, T., Bakó, I., Pápai, I. & Horváth, A. (2003). *J. Phys. Chem. A*, **107**, 9903–9909.
- Moret, M.-E., Tavernelli, I. & Rothlisberger, U. (2009). *J. Phys. Chem. B*, **113**, 7737–7744.
- Mulfort, K. L., Mukherjee, A., Kokhan, O., Du, P. & Tiede, D. M. (2013). *Chem. Soc. Rev.* **42**, 2215–2227.
- Mustafi, D. & Makinen, M. W. (1988). *Inorg. Chem.* **27**, 3360–3368.
- Neder, R. B. & Proffen, Th. (2020). *J. Appl. Cryst.* **53**, 710–721.
- Neuefeind, J., Skanthakumar, S. & Soderholm, L. (2004). *Inorg. Chem.* **43**, 2422–2426.
- Park, S., Bardhan, J. P., Roux, B. & Makowski, L. (2009). *J. Chem. Phys.* **130**, 134114.
- Penfold, T. J., Reinhard, M., Rittmann-Frank, M. H., Tavernelli, I., Rothlisberger, U., Milne, C. J., Glatzel, P. & Chergui, M. (2014). *J. Phys. Chem. A*, **118**, 9411–9418.
- Perera, P., Wyche, M., Loethen, Y. & Ben-Amotz, D. (2008). *J. Am. Chem. Soc.* **130**, 4576–4577.
- Peterson, P. F., Olds, D., McDonnell, M. T. & Page, K. (2021). *J. Appl. Cryst.* **54**, 317–332.
- Qiu, X., Thompson, J. W. & Billinge, S. J. L. (2004). *J. Appl. Cryst.* **37**, 678–678.
- Rajput, N. N., Seguin, T. J., Wood, B. M., Qu, X. & Persson, K. A. (2018). *Top. Curr. Chem. (Z.)*, **376**, 19.
- Rappe, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A. I. & Skiff, W. M. (1992). *J. Am. Chem. Soc.* **114**, 10024–10035.
- Roelsgaard, M., Dippel, A.-C., Borup, K. A., Nielsen, I. G., Broge, N. L. N., Röh, J. T., Gutowski, O. & Iversen, B. B. (2019). *IUCrJ*, **6**, 299–304.
- Scalambra, F., Imberti, S., Holzmann, N., Bernasconi, L. & Romerosa, A. (2021). *ACS Omega*, **6**, 1751–1757.
- Sengupta, A., Li, Z., Song, L. F., Li, P. & Merz, K. M. J. (2021). *J. Chem. Inf. Model.* **61**, 869–880.
- Singh, U. C. & Kollman, P. A. (1984). *J. Comput. Chem.* **5**, 129–145.
- Skanthakumar, S., Jin, G. B., Lin, J., Vallet, V. & Soderholm, L. (2017). *J. Phys. Chem. B*, **121**, 8577–8584.
- Skinner, L. B., Huang, C., Schlesinger, D., Pettersson, L. G. M., Nilsson, A. & Benmore, C. J. (2013). *J. Chem. Phys.* **138**, 074506.
- Soderholm, L., Skanthakumar, S. & Neuefeind, J. (2005). *Anal. Bioanal. Chem.* **383**, 48–55.
- Soper, A. K. (1997). *J. Phys. Condens. Matter*, **9**, 2399–2410.
- Staroverov, V. N., Scuseria, G. E., Tao, J. & Perdew, J. P. (2003). *J. Chem. Phys.* **119**, 12129–12137.
- Sun, Y. & Petersen, P. B. (2017). *J. Phys. Chem. Lett.* **8**, 611–614.
- Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.
- Terban, M. W. & Billinge, S. J. L. (2022). *Chem. Rev.* **122**, 1208–1272.
- Thijssse, B. J. (1984). *J. Appl. Cryst.* **17**, 61–76.
- Tiede, D. M., Mardis, K. L. & Zuo, X. (2009). *Photosynth. Res.* **102**, 267–279.
- Tiede, D. M., Zhang, R., Chen, L. X., Yu, L. & Lindsey, J. S. (2004). *J. Am. Chem. Soc.* **126**, 14054–14062.
- Toby, B. H. & Von Dreele, R. B. (2013). *J. Appl. Cryst.* **46**, 544–549.
- Waasmaier, D. & Kirfel, A. (1995). *Acta Cryst.* **A51**, 416–431.

- Warren, B. E. (1990). *X-ray Diffraction*. New York: Dover Publications.
- Weigend, F. & Ahlrichs, R. (2005). *Phys. Chem. Chem. Phys.* **7**, 3297–3305.
- Weng, J., Xu, W., Wiaderek, K. M., Borkiewicz, O. J., Chen, J., Von Dreele, R. B., Gallington, L. C. & Ruett, U. (2023). *J. Synchrotron Rad.* **30**, 546–554.
- Xie, Z.-L., Liu, X., Valentine, A. J. S., Lynch, V. M., Tiede, D. M., Li, X. & Mulfort, K. L. (2022). *Angew. Chem. Int. Ed.* **61**, e202111764.
- Yang, K. R., Matula, A. J., Kwon, G., Hong, J., Sheehan, S. W., Thomsen, J. M., Brudvig, G. W., Crabtree, R. H., Tiede, D. M., Chen, L. X. & Batista, V. S. (2016). *J. Am. Chem. Soc.* **138**, 5511–5514.