



Edited by S. S. Hasnain, University of Liverpool, England

Keywords: diffuse scattering; intermolecular correlations; LLM models.

Interactions that know no boundaries

Michael E. Wall*

Computer, Computational and Statistical Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA. *Correspondence e-mail: mewall@lanl.gov

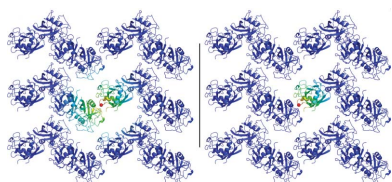
Deviations from an ideal crystal lead to diffuse scattering (DS) intensity, both between and beneath the Bragg peaks in diffraction patterns (Guinier, 1963). First characterized using simple ionic crystals in early studies of X-ray diffraction (Lonsdale, 1942), DS has a rich history (Welberry & Weber, 2016) and is a well established technique in small-molecule crystallography (Welberry, 2004). DS studies in macromolecular crystallography began more recently (Phillips *et al.*, 1980) and now the potential for obtaining information about protein motions is fueling the growing interest in DS (Meisburger *et al.*, 2017).

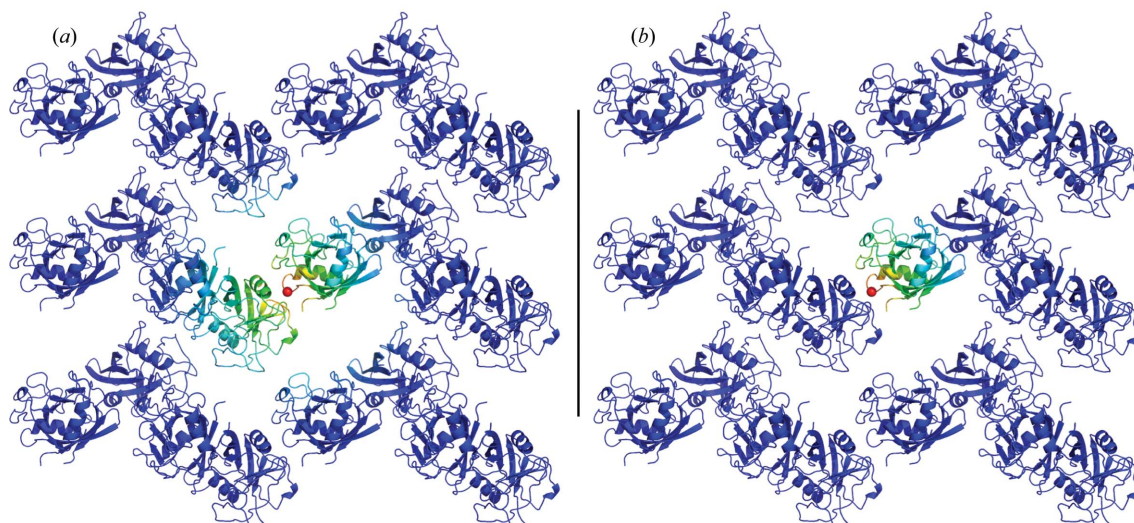
As noted in a previous **IUCrJ** commentary (Keen, 2016), accurate modeling of small-molecule DS requires not only information about the variations of individual molecules or unit cells, but also information about the correlated variations in a more extended environment. Similarly, macromolecular DS studies indicate the importance of modeling interactions across unit-cell boundaries in normal-modes models (Riccardi *et al.*, 2010), as well as the molecular dynamics models (Wall, 2018) of macromolecular diffuse scattering that are shown in this issue. The liquid-like motions (LLM) model (Caspar *et al.*, 1988), in which the correlated variations are modeled as if the crystal were a soft homogeneous material, explains the overall DS pattern in several protein crystals (Caspar *et al.*, 1988; Clarage *et al.*, 1992; Van Benschoten *et al.*, 2016; Wall, Clarage & Phillips, 1997; Wall, Ealick & Gruner, 1997). However, the consequences of including intermolecular interactions for the accuracy of the LLM model were not clear until now.

In this issue, Peck and co-workers (Peck *et al.*, 2018) investigate the importance of intermolecular interactions by assessing the accuracy of two alternative versions of an LLM model (Caspar *et al.*, 1988) (Fig. 1). In the original version of the model, the correlations extend across molecular boundaries (Fig. 1*a*). In this case, the diffuse intensity is derived from the squared crystal transform, which is sharply peaked. In a modified version of the model, correlations terminate at the molecular boundary (Fig. 1*b*). In this case, the diffuse intensity is derived from the squared molecular transform of the asymmetric unit (in the cases considered, a single molecule), which is continuous in reciprocal space. In both cases, the transform is blurred; shorter correlation lengths correspond to a larger scale blurring of the transform. Both models are optimized to maximize the agreement with the data, enabling a well controlled comparison.

To be consistent with the state of the art (Meisburger *et al.*, 2017), three-dimensional diffuse datasets were used for the comparison, obtained from crystalline cyclophilin A (CypA) (PDB entry 4yuo; Fraser, 2015), WrpA (PDB entry 5f51; Herrou & Crosson, 2015) and alkaline phosphatase (PDB entry 5c66; Peck *et al.*, 2017). The CypA data were the subject of a prior DS study (Van Benschoten *et al.*, 2016) and the others were newly analyzed for this study, providing valuable additions to the currently limited amount of available macromolecular DS data. For all three datasets, the original LLM model, which includes intermolecular correlations, was substantially more accurate than the modified model in which molecules are independent. Instances of elastic network models (ENMs) (Bray *et al.*, 2011) and ensemble models (van den Bedem *et al.*, 2009) were tested, but since these models were only narrowly explored and did not incorporate prior insight into their use for diffuse scattering calculations (Levin *et al.*, 2007; Riccardi *et al.*, 2010), the tests were inconclusive. Models of independent rigid-body motions (Moore, 2009) also were compared; their accuracy was similar to that of the independent LLM model, providing additional evidence for the importance of intermolecular interactions in macromolecular DS.

As illustrated in this issue (Peck *et al.*, 2018; Wall, 2018), for a small number of cases, DS studies provide insight into what is really going on in macromolecular crystals, beyond what can be discerned using Bragg analysis. However, DS data collection and processing




Figure 1

Alternative models of correlated displacements in a liquid-like motions model, compared by Peck *et al.* (2018). Each panel depicts a 2×3 unit-cell section extracted from the (010)–(001) layer of crystalline cyclophilin A [PDB entry 2cpl (Ke, 1992), used in Peck *et al.* (2018)]. The *b* axis is aligned with the vertical direction, and the *c* axis with the horizontal. Each $P2_12_1$ unit cell contains four copies of the protein that arrange into a sawtooth in this projection. The strength of correlation with an atom near the center (Asp26 Ca, highlighted as a red sphere) is indicated using a heat map. The correlation decays exponentially with distance, according to a liquid-like motions model (Caspar *et al.*, 1988). (a) The original model, in which correlations extend to atoms on different proteins. (b) A modified model, in which correlations are limited to atoms on the same protein. Peck and coworkers (Peck *et al.*, 2018) found that diffuse scattering data for this and two other systems favor model (a) over model (b). The figure was created using *PyMol* (<https://sourceforge.net/projects/pymol/>).

is less well developed in comparison with the Bragg analysis, and model accuracy is still lacking. Until general insights are available from a larger number of cases, it would be wise not to dismiss any individual approach to analyzing the data. As for the Bragg data, each DS data point is influenced by the whole system; therefore we can expect that the entire DS model will need to be sufficiently accurate before the atomic details of macromolecular movements can be revealed.

Acknowledgements

This is Los Alamos National Laboratory technical release LA-UR-18-20721.

Funding information

Support was provided by the US Department of Energy via the Exascale Computing Project and under contract DE-AC52-06NA25396 to Los Alamos National Security, LLC. Additional support was provided by the University of California via the Laboratory Fees Research Program.

References

- Bedem, H. van den, Dhanik, A., Latombe, J.-C. & Deacon, A. M. (2009). *Acta Cryst.* **D65**, 1107–1117.
- Bray, J. K., Weiss, D. R. & Levitt, M. (2011). *Biophys. J.* **101**, 2966–2969.
- Caspar, D. L., Clarage, J., Salunke, D. M. & Clarage, M. (1988). *Nature*, **332**, 659–662.
- Clarage, J. B., Clarage, M. S., Phillips, W. C., Sweet, R. M. & Caspar, D. L. (1992). *Proteins*, **12**, 145–157.
- Fraser, J. S. (2015). *X-ray diffraction data for Cyclophilin A*. PDB code 4YUO. SBCGrid Data Bank v.1. <https://doi.org/10.15785/sbgrid/68>.
- Guinier, A. (1963). *X-ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies*. San Francisco: W. H. Freeman and Co.
- Herrou, J. & Crosson, S. (2015). *X-ray diffraction data for B. abortus WrbA related protein A (apo-WrpA)*. PDB code 5F51. SBCGrid Data Bank v.1. <https://doi.org/10.15785/sbgrid/203>.
- Ke, H. (1992). *J. Mol. Biol.* **228**, 539–550.
- Keen, D. A. (2016). *IUCrJ*, **3**, 8–9.
- Levin, E. J., Kondrashov, D. A., Wesenberg, G. E. & Phillips, G. N. Jr (2007). *Structure*, **15**, 1040–1052.
- Lonsdale, K. (1942). *Proc. Phys. Soc.* **54**, 314–353.
- Meisburger, S. P., Thomas, W. C., Watkins, M. B. & Ando, N. (2017). *Chem. Rev.* **117**, 7615–7672.
- Moore, P. B. (2009). *Structure*, **17**, 1307–1315.
- Peck, A., Poitevin, F. & Lane, T. J. (2018). *IUCrJ*, **5**, 211–222.
- Peck, A., Ressler, S. & Herschlag, D. (2017). *X-ray diffraction data for E. coli alkaline phosphatase in complex with tungstate*. PDB code 5C66. SBCGrid Data Bank v.1. <https://doi.org/10.15785/sbgrid/456>.
- Phillips, G. N. Jr, Fillers, J. P. & Cohen, C. (1980). *Biophys. J.* **32**, 485–502.
- Riccardi, D., Cui, Q. & Phillips, G. N. Jr (2010). *Biophys. J.* **99**, 2616–2625.
- Van Benschoten, A. H., Liu, L., Gonzalez, A., Brewster, A. S., Sauter, N. K., Fraser, J. S. & Wall, M. E. (2016). *Proc. Natl Acad. Sci. USA*, **113**, 4069–4074.
- Wall, M. E. (2018). *IUCrJ*, **5**, 172–181.
- Wall, M. E., Clarage, J. B. & Phillips, G. N. (1997). *Structure*, **5**, 1599–1612.
- Wall, M. E., Ealick, S. E. & Gruner, S. M. (1997). *Proc. Natl Acad. Sci. USA*, **94**, 6180–6184.
- Welberry, T. R. (2004). *Diffuse X-ray Scattering and Models of Disorder*. Oxford University Press.
- Welberry, T. R. & Weber, T. (2016). *Crystallogr. Rev.* **22**, 2–78.