Structure
Meeting Review

Diffuse X-Ray Scattering to Model Protein Motions

Michael E. Wall,[1,*] Paul D. Adams,[2,3] James S. Fraser,[4] and Nicholas K. Sauter[2]
[1]Computer, Computational, and Statistical Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[2]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[3]Department of Bioengineering, University of California, Berkeley, Berkeley, CA 94720, USA
[4]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA 94143, USA
*Correspondence: mewall@lanl.gov
http://dx.doi.org/10.1016/j.str.2014.01.002

Problems in biology increasingly need models of protein flexibility to understand and control protein function. At the same time, as they improve, crystallographic methods are marching closer to the limits of what can be learned from Bragg data in isolation. It is thus inevitable that mainstream protein crystallography will turn to diffuse scattering to model protein motions and improve crystallographic models. The time is ripe to make it happen.

Researchers recently gathered in Berkeley, California to assess the state of the art diffuse X-ray scattering and the potential for using it as an experimental probe of protein motions. The workshop entitled ''Can Diffuse X-Ray Scattering Reveal Protein Dynamics''? was hosted by the Advanced Light Source User Meeting on October 9, 2013. It consisted of nine presentations that addressed topics such as data collection and integration, modeling and simulation, and dissemination of data and methods. This meeting report summarizes our current perspective on the field, informed by the deliberations of the workshop.

In traditional protein crystallography, the Bragg peak measurements are used to derive the mean unit cell electron density, which is used for building and validating single structural models. By contrast, the diffuse scattering, which lies between (and under) the Bragg peaks, contains rich information about the two-point correlations of the electron density fluctuations (Amorós and Amorós, 1968; Guinier, 1963; James, 1948; Warren, 1969; Welberry, 2004; Willis and Pryor, 1975; Wooster, 1962; Zachariasen, 1945). Diffuse scattering is potentially a powerful constraint for modeling protein motions, because, unlike the Bragg peaks, it contains information about which atoms move together.

Although diffuse scattering is an established technique in materials science, its use in protein crystallography so far has been confined to a relatively small number of pioneering studies (Caspar et al., 1988; Chacko and Phillips, 1992; Clarage et al., 1992, 1995; Doucet and Benoit, 1987; Glover et al., 1991; Héry et al., 1998; Kolatkar et al., 1994; Meinhold et al., 2007; Meinhold and Smith, 2005a, 2005b, 2007; Mizuguchi et al., 1994; Moore, 2009; Phillips et al., 1980; Riccardi et al., 2010; Wall et al., 1997a, 1997b). The motivation for increasing the use of diffuse scattering in protein crystallography is now strong. Crystallography is producing increasingly sophisticated and detailed models of protein motions such as translation-libration-screw (TLS) models (Chaudhry et al., 2004) and contact network models (van den Bedem et al., 2013). Whereas the Bragg data cannot distinguish among different models that yield similar mean electron density, diffuse scattering can distinguish models by their correlated motions. Diffuse scattering therefore might be developed into a powerful tool for modeling crystalline protein motions. Indeed, in *Structure*, Peter Moore (Moore, 2009) has argued that diffuse scattering should be used to test TLS models, and Mark Wilson (Wilson, 2013) has noted that diffuse scattering could be used to develop and validate more detailed contact network (van den Bedem et al., 2013) or ensemble models (Burnley et al., 2012).

Increasing the use of diffuse scattering in protein crystallography naturally presents some challenges. These challenges can be overcome (and should be met) by appropriate efforts. Progress is needed on several fronts.

Data Collection

Data collection procedures and equipment for precise measurement of diffuse scattering must become widespread in protein crystallography. The considerations are similar to those for small-angle scattering, which places similar emphasis on the elimination of systematic background noise. Previously successful procedures have been documented and are compatible with traditional crystallography experiments (Wall, 2009). Charge-coupled-device (CCD) detectors at synchrotrons (Walter et al., 1995) have been successfully used to collect full three-dimensional diffuse scattering data sets (Wall et al., 1997b). The negligibly small point spread function and higher dynamic range of pixel array detectors (PADs) (Gruner, 2012) are now enabling improved data collection compared to CCDs, especially to resolve fine scale features in the neighborhood of Bragg peaks. Simultaneous Bragg and diffuse scattering data already have been collected from protein crystals using a PAD-based PILATUS detector (Andrew VanBenschoten and J.S.F., unpublished data), and the meeting participants noted the importance of the temperature dependence of diffuse features as an important future direction.

Data Integration

Methods for diffuse scattering data integration must be extended and combined with Bragg integration in the standard crystallography toolkit. Two complementary approaches can be exploited. One approach, adopted by Lunus software (Wall, 2009; http://lunus.sf.net), is to collect diffuse scattering measurements on a three-dimensional reciprocal space lattice, the structure of which is identical to a Bragg lattice. An advantage of this approach is that it enables the existing infrastructure in

crystallography to be maximally leveraged, from data indexing (a preliminary pipeline using LABELIT [Sauter et al., 2004] and CCTBX [Grosse-Kunstleve et al., 2002] was demonstrated at the workshop) to model building, validation, refinement, and tools for computing model inputs such as structure factors (such as are available in the PHENIX software suite [Adams et al., 2010]; http://www.phenix-online.org/). It also can be adapted to measure small-scale, streaked diffuse features by constructing a lattice with a finer sampling of reciprocal space (Wall et al., 1997a). A challenge shared with traditional methods is the need to separate the diffuse from the Bragg intensity; in this sense, better modeling of the diffuse scattering can potentially improve Bragg peak integration. A second approach is to use the diffraction images themselves to constrain models of the protein crystal. This approach is inspired by the EVAL15 method (Schreurs et al., 2010; http://www.crystal.chem.uu.nl/distr/eval/) in which a physical model of the crystal is refined using highly detailed Bragg peak profiles obtained from diffraction images. An advantage of this approach is that it enables model building using the primary data and eliminates the need for separating the Bragg and diffuse intensity. A challenge is the additional computational cost, which can potentially be overcome by taking advantage of GPUs or other advanced computing architectures.

Model Building and Refinement

New model building and refinement tools must be developed for diffuse scattering. Many diffuse scattering studies have addressed the forward problem of calculating a simulated diffraction image or three-dimensional diffuse lattice from a model of correlated motions and comparing it to experimental data. The forward problem has been a useful paradigm for validating correlated motions predicted from molecular dynamics (MD) simulations (Clarage et al., 1995; Doucet and Benoit, 1987; Faure et al., 1994; Héry et al., 1998; Meinhold and Smith, 2005a, 2007) and normal mode analysis (Meinhold et al., 2007; Mizuguchi et al., 1994; Riccardi et al., 2010), and we expect it to be useful for validating predicted correlations from TLS models, contact networks, and other independently developed models. To fully realize its potential, however, diffuse scattering should also be developed as a tool for solving the inverse problem of deriving models of correlated motions from the data. One important advance would be to derive atom pair-coupled displacement parameters (i.e., the off-diagonal elements of the covariance matrix of atomic displacements) from diffuse scattering data. Initial steps have been taken by using individual diffraction images (Caspar et al., 1988; Chacko and Phillips, 1992; Clarage et al., 1992) or three-dimensional diffuse scattering data (Wall et al., 1997a, 1997b) to model homogeneous elastic or liquid-like correlated motions in the crystal.

Molecular Dynamics Simulations

One especially important challenge in model building is the derivation of ensemble models from MD simulations. This approach has been revisited over the years as advances in computers and algorithms have enabled longer simulations of larger systems and deeper sampling of conformational ensembles. Early MD simulations of single molecules of lysozyme (Faure et al., 1994) and myoglobin (Clarage et al., 1995) yielded ensemble models that did not accurately reproduce diffuse scattering data. Clarage et al. (1995) traced the problem to inadequate sampling of the conformational ensemble, leading to a lack of convergence of the covariance matrix of atomic displacements (the lysozyme simulation was 600 ps duration and myoglobin was 500 ps). Later, 1 ns simulations of a single $P2_12_12_1$ lysozyme unit cell (Héry et al., 1998) and 10 ns simulations of a single $P4_1$ staphylococcal nuclease unit cell (Meinhold and Smith, 2005b) showed improved agreement with the data but did not show complete convergence. Longer MD simulations of staphylococcal nuclease (M.E.W., unpublished data) indicate that convergence is now within reach even for models consisting of multiple unit cells, which might be important to accurately describe diffuse scattering data. Future avenues for achieving increased sampling include millisecond all-atom simulations (Dror et al., 2012), advanced sampling methods such as parallel tempering (Earl and Deem, 2005), and acceleration schemes such as Markov State Models (Pande et al., 2010).

Other challenges remain. For example, there are only a limited number of data sets available. This is now changing, with several groups ramping up their data collection efforts. Dissemination of data would be facilitated by establishing a resource where existing diffuse scattering data and models could be archived and made publicly available using standard formats. We are seeking advice from the Worldwide Protein Data Bank as how to best manage these data, including the development of data definitions for the PDBx/mmCIF dictionary.

Problems in biology increasingly need models of protein flexibility to understand and control protein function. At the same time, as they improve, crystallographic methods are marching closer to the limits of what can be learned from Bragg data in isolation. It is thus inevitable that mainstream protein crystallography will turn to diffuse scattering to model protein motions and improve crystallographic models. The time is ripe to make it happen.

ACKNOWLEDGMENTS

We are grateful to the participants of the workshop, both for their insights into diffuse scattering and for a highly stimulating meeting. M.E.W. is supported by the Department of Energy through the Laboratory-Directed Research and Development Program at Los Alamos National Laboratory; P.D.A. is supported by National Institutes of Health grant 1P01GM063210; J.S.F. is supported by a National Institutes of Health Early Independence Award (DP5OD009180) and QB3; and N.K.S. is supported by National Institutes of Health grant GM095887. The Advanced Light Source User Meeting was supported by the Department of Energy under Contract DE-AC02-05CH11231.

The following participants gave talks at the workshop: Gregory R. Bowman, University of California, Berkeley; James Holton, Lawrence Berkeley National Laboratory (LBNL); George N. Phillips, Jr., Rice University; Demian Riccardi, Oak Ridge National Laboratory (ORNL); N.K.S., LBNL; David E. Skinner, National Energy Research Scientific Computing Center; Jeremy C. Smith, ORNL; Andrew VanBenschoten, University of California, San Francisco; and M.E.W., Los Alamos National Laboratory. Slides from the workshop are available at http://cci.lbl.gov/dials/oct_2013_diffuse.htm.

REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. *66*, 213–221.

Amorós, J.L., and Amorós, M. (1968). Molecular Crystals; Their Transforms and Diffuse Scattering. (New York: Wiley).

Burnley, B.T., Afonine, P.V., Adams, P.D., and Gros, P. (2012). Modelling dynamics in protein crystal structures by ensemble refinement. eLife Sciences *1*, e00311.

Caspar, D.L., Clarage, J., Salunke, D.M., and Clarage, M. (1988). Liquid-like movements in crystalline insulin. Nature *332*, 659–662.

Chacko, S., and Phillips, G.N., Jr. (1992). Diffuse x-ray scattering from tropomyosin crystals. Biophys. J. *61*, 1256–1266.

Chaudhry, C., Horwich, A.L., Brunger, A.T., and Adams, P.D. (2004). Exploring the structural dynamics of the *E.coli* chaperonin GroEL using translation-libration-screw crystallographic refinement of intermediate states. J. Mol. Biol. *342*, 229–245.

Clarage, J.B., Clarage, M.S., Phillips, W.C., Sweet, R.M., and Caspar, D.L. (1992). Correlations of atomic movements in lysozyme crystals. Proteins *12*, 145–157.

Clarage, J.B., Romo, T., Andrews, B.K., Pettitt, B.M., and Phillips, G.N., Jr. (1995). A sampling problem in molecular dynamics simulations of macromolecules. Proc. Natl. Acad. Sci. USA *92*, 3288–3292.

Doucet, J., and Benoit, J.P. (1987). Molecular dynamics studied by analysis of the X-ray diffuse scattering from lysozyme crystals. Nature *325*, 643–646.

Dror, R.O., Dirks, R.M., Grossman, J.P., Xu, H., and Shaw, D.E. (2012). Biomolecular simulation: a computational microscope for molecular biology. Annu Rev Biophys *41*, 429–452.

Earl, D.J., and Deem, M.W. (2005). Parallel tempering: theory, applications, and new perspectives. Phys. Chem. Chem. Phys. *7*, 3910–3916.

Faure, P., Micu, A., Pérahia, D., Doucet, J., Smith, J.C., and Benoit, J.P. (1994). Correlated intramolecular motions and diffuse X-ray scattering in lysozyme. Nat. Struct. Biol. *1*, 124–128.

Glover, I.D., Harris, G.W., Helliwell, J.R., and Moss, D.S. (1991). The variety of X-ray diffuse scattering from macromolecular crystals and its respective components. Acta Cryst *B47*, 960–968.

Grosse-Kunstleve, R.W., Sauter, N.K., Moriarty, N.W., and Adams, P.D. (2002). The Computational Crystallography Toolbox: crystallographic algorithms in a reusable software framework. J. Appl. Cryst. *35*, 126–136.

Gruner, S.M. (2012). Imaging X-ray detectors open new vistas. Phys. Today *65*, 29–34.

Guinier, A. (1963). X-ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies. (San Francisco: W. H. Freeman and Company).

Héry, S., Genest, D., and Smith, J.C. (1998). X-ray diffuse scattering and rigid-body motion in crystalline lysozyme probed by molecular dynamics simulation. J. Mol. Biol. *279*, 303–319.

James, R. (1948). The Optical Principles of the Diffraction of X-Rays. (London: Bell).

Kolatkar, A.R., Clarage, J.B., and Phillips, G.N., Jr. (1994). Analysis of diffuse scattering from yeast initiator tRNA crystals. Acta Crystallogr. D Biol. Crystallogr. *50*, 210–218.

Meinhold, L., and Smith, J.C. (2005a). Correlated dynamics determining x-ray diffuse scattering from a crystalline protein revealed by molecular dynamics simulation. Phys. Rev. Lett. *95*, 218103.

Meinhold, L., and Smith, J.C. (2005b). Fluctuations and correlations in crystalline protein dynamics: a simulation analysis of staphylococcal nuclease. Biophys. J. *88*, 2554–2563.

Meinhold, L., and Smith, J.C. (2007). Protein dynamics from X-ray crystallography: anisotropic, global motion in diffuse scattering patterns. Proteins *66*, 941–953.

Meinhold, L., Merzel, F., and Smith, J.C. (2007). Lattice dynamics of a protein crystal. Phys. Rev. Lett. *99*, 138101.

Mizuguchi, K., Kidera, A., and Gō, N. (1994). Collective motions in proteins investigated by X-ray diffuse scattering. Proteins *18*, 34–48.

Moore, P.B. (2009). On the relationship between diffraction patterns and motions in macromolecular crystals. Structure *17*, 1307–1315.

Pande, V.S., Beauchamp, K., and Bowman, G.R. (2010). Everything you wanted to know about Markov State Models but were afraid to ask. Methods *52*, 99–105.

Phillips, G.N., Jr., Fillers, J.P., and Cohen, C. (1980). Motions of tropomyosin. Crystal as metaphor. Biophys. J. *32*, 485–502.

Riccardi, D., Cui, Q., and Phillips, G.N., Jr. (2010). Evaluating elastic network models of crystalline biological molecules with temperature factors, correlated motions, and diffuse x-ray scattering. Biophys. J. *99*, 2616–2625.

Sauter, N.K., Grosse-Kunstleve, R.W., and Adams, P.D. (2004). Robust indexing for automatic data collection. J. Appl. Cryst. *37*, 399–409.

Schreurs, A.M.M., Xian, X., and Kroon-Batenburg, L.M.J. (2010). EVAL15: a diffraction data integration method based on ab initio predicted profiles. J. Appl. Cryst. *43*, 70–82.

van den Bedem, H., Bhabha, G., Yang, K., Wright, P.E., and Fraser, J.S. (2013). Automated identification of functional dynamic contact networks from X-ray crystallography. Nat. Methods *10*, 896–902.

Wall, M.E. (2009). Methods and software for diffuse X-ray scattering from protein crystals. Methods Mol. Biol. *544*, 269–279.

Wall, M.E., Clarage, J.B., and Phillips, G.N. (1997a). Motions of calmodulin characterized using both Bragg and diffuse X-ray scattering. Structure *5*, 1599–1612.

Wall, M.E., Ealick, S.E., and Gruner, S.M. (1997b). Three-dimensional diffuse x-ray scattering from crystals of staphylococcal nuclease. Proc. Natl. Acad. Sci. USA *94*, 6180–6184.

Walter, R.L., Thiel, D.J., Barna, S.L., Tate, M.W., Wall, M.E., Eikenberry, E.F., Gruner, S.M., and Ealick, S.E. (1995). High-resolution macromolecular structure determination using CCD detectors and synchrotron radiation. Structure *3*, 835–844.

Warren, B.E. (1969). X-Ray Diffraction. (Reading, MA: Addison-Wesley).

Welberry, T.R. (2004). Diffuse X-Ray Scattering and Models of Disorder. (Oxford: Oxford University Press).

Willis, B.T.M., and Pryor, A.W. (1975). Thermal Vibrations in Crystallography. (Cambridge: Cambridge University Press).

Wilson, M.A. (2013). Visualizing networks of mobility in proteins. Nat. Methods *10*, 835–837.

Wooster, W.A. (1962). Diffuse X-Ray Reflections from Crystals. (Oxford: Oxford University Press).

Zachariasen, W. (1945). Theory of X-Ray Diffraction in Crystals. (New York: Wiley).

 ©2014 Elsevier Ltd All rights reserved