Metadata needed for the full exploitation of diffuse scattering data from protein crystals

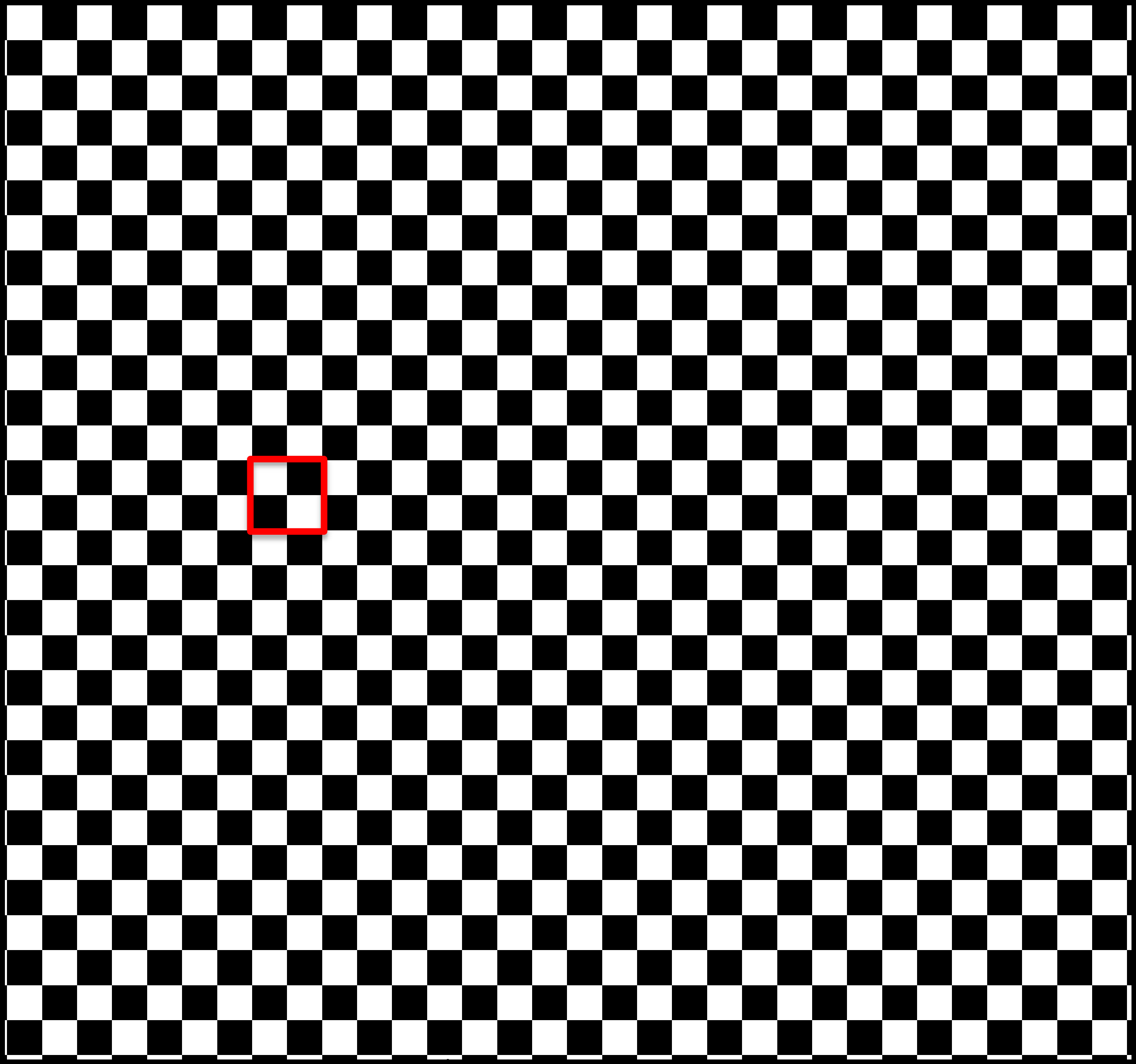
Michael Wall

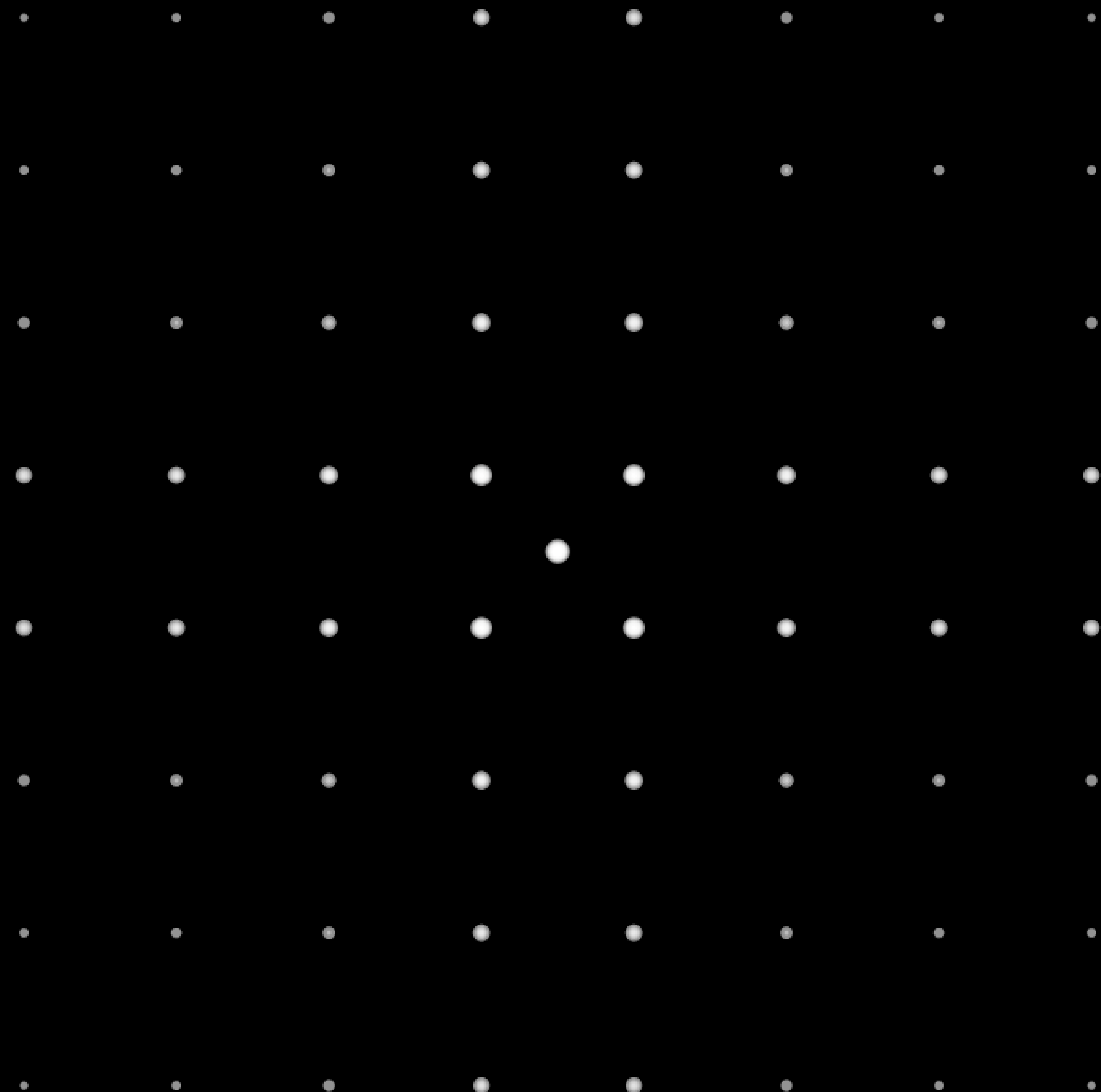
Los Alamos National Laboratory

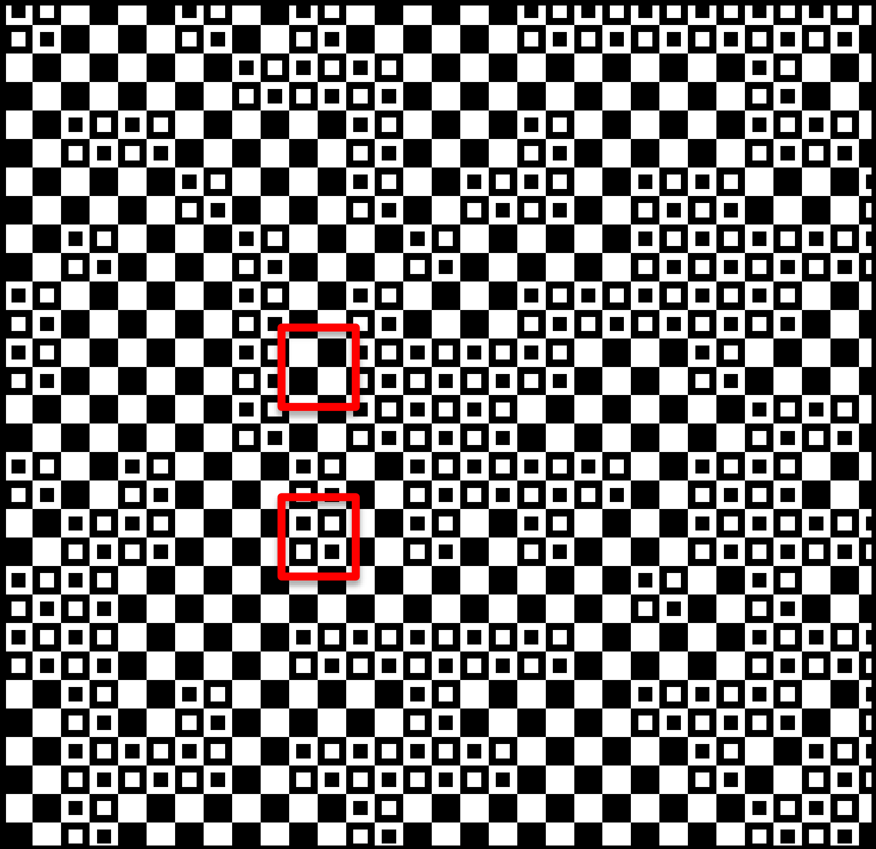
ECM 29 Satellite Workshop on Metadata for Raw Data from X-ray Diffraction and Other Structural Techniques

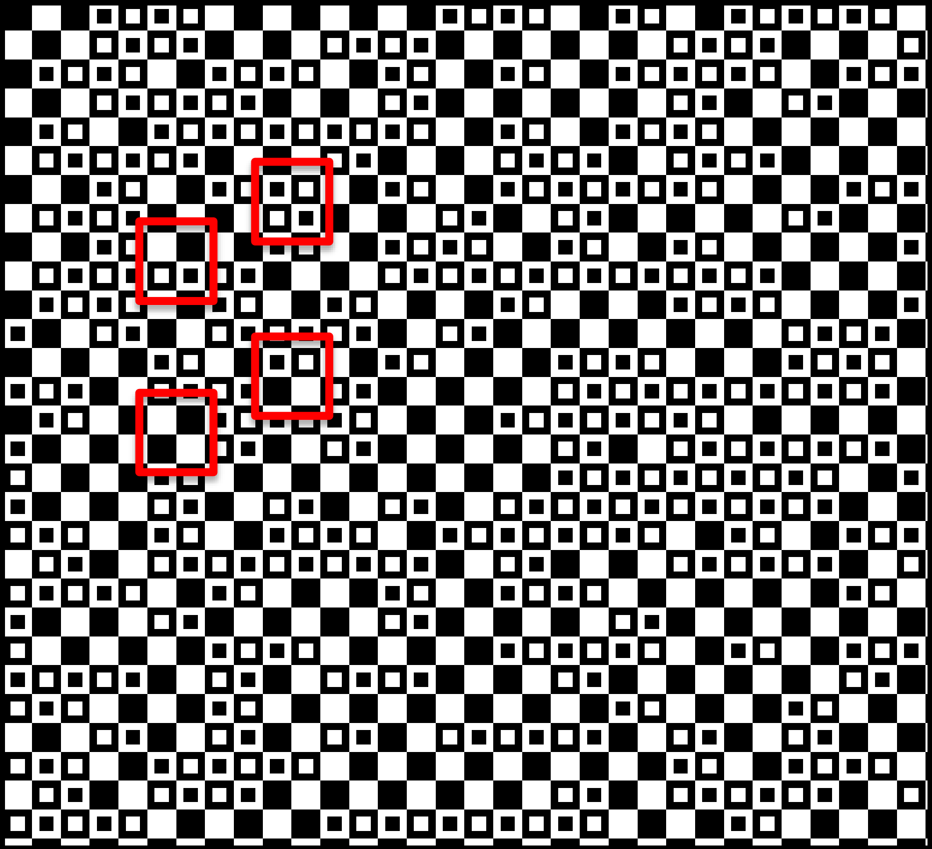
Rovinj, Croatia

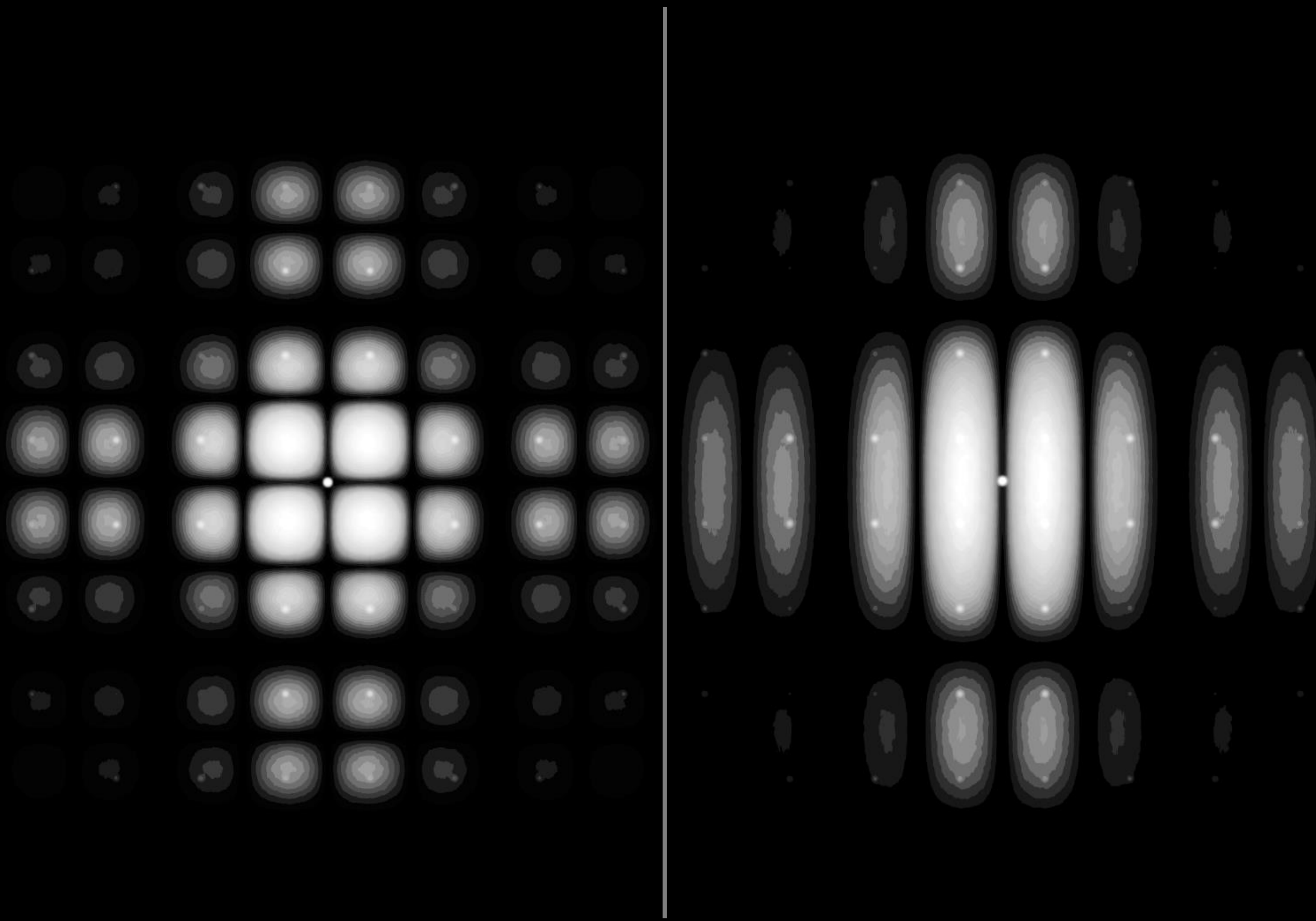
22-23 August 2015

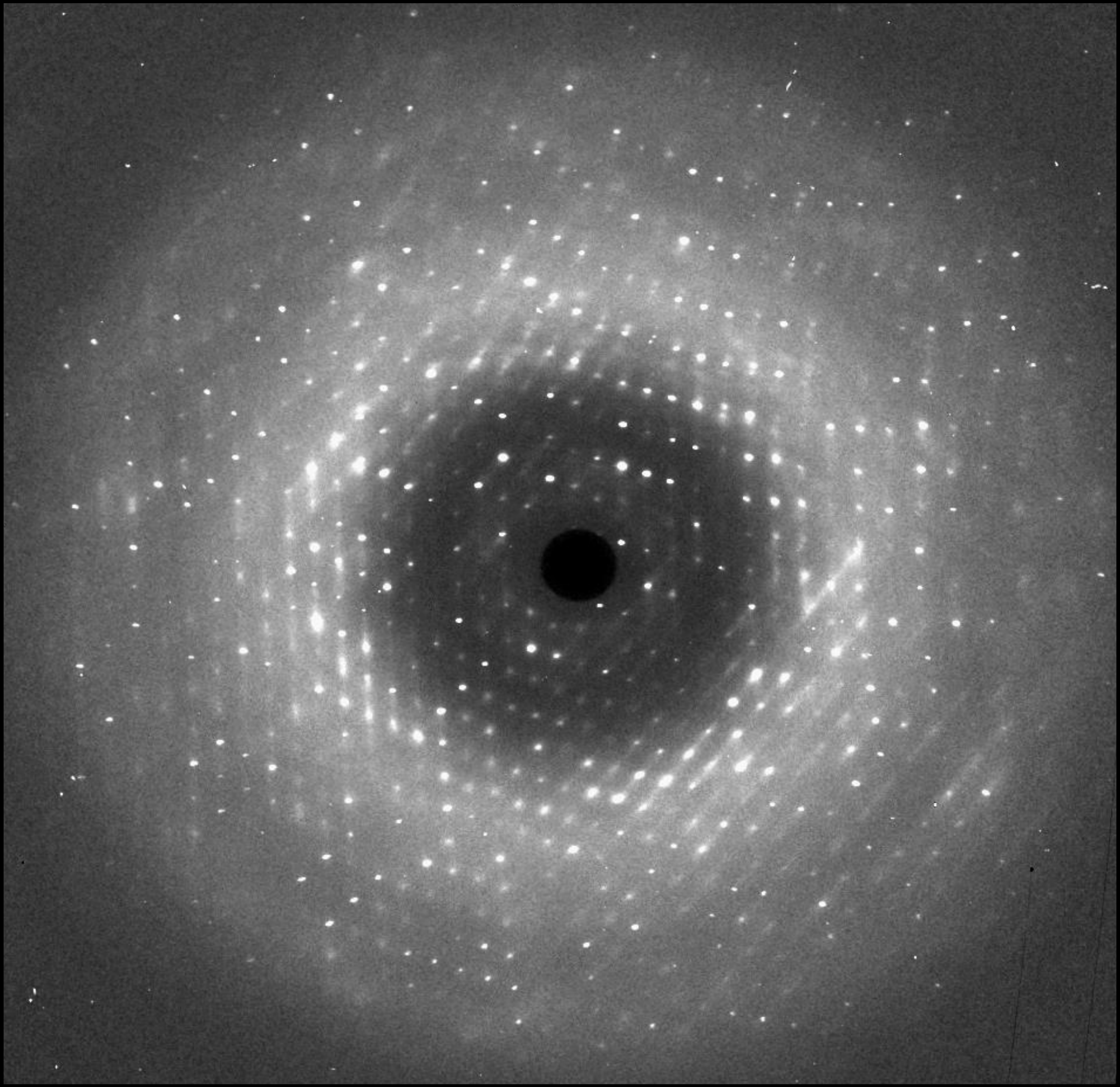












Diffuse X-Ray Scattering to Model Protein Motions

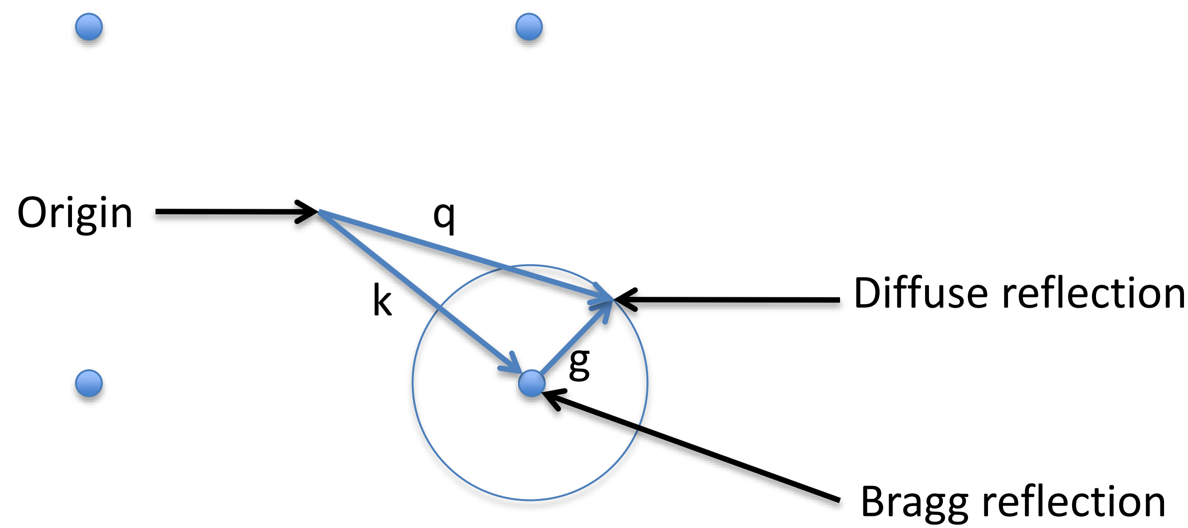
Michael E. Wall,[1,*] Paul D. Adams,[2,3] James S. Fraser,[4] and Nicholas K. Sauter[2]
[1]Computer, Computational, and Statistical Sciences Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
[2]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, 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.

Structure **22** (2014) 182

Kathleen Lonsdale
Diffuse Scattering Pioneer

- 1924, student of Lawrence Bragg
- 1928, Solved benzene structure
  - Ended 60 year debate about flat aromatic ring
- 1942, **Champion of diffuse X-ray scattering**
- 1945, Fellow of Royal Society
  - One of first two women (along with Marjory Stephenson)
- 1956, Dame Commander of OBE
- 1966, First woman president of IUCr



$$I_D(\mathbf{q}) = \sum_{\mathbf{g},s} \left[I_0(\mathbf{q}+\mathbf{g}) + I_0(\mathbf{q}-\mathbf{g}) \right] \left| \mathbf{q} \cdot \mathbf{v}_s(\mathbf{g}) \right|^2$$

Born & Sarginson, 1941
James, 1948

Thermal Diffuse Scattering (TDS)

$$F = \sum_n f_n e^{i\mathbf{q}\cdot\mathbf{R}_n}$$

$$I(t) = |F|^2 = \sum_n \sum_m f_n f_m^* e^{i\mathbf{q}\cdot(\mathbf{R}_n - \mathbf{R}_m)}$$

$$I = \sum_n \sum_m \left[\left(\left\langle |f_n|^2 \right\rangle - |\langle f_n \rangle|^2 \right) \delta_{nm} + |\langle f_n \rangle|^2 \right] e^{i\mathbf{q}\cdot(\mathbf{R}_n - \mathbf{R}_m)}$$

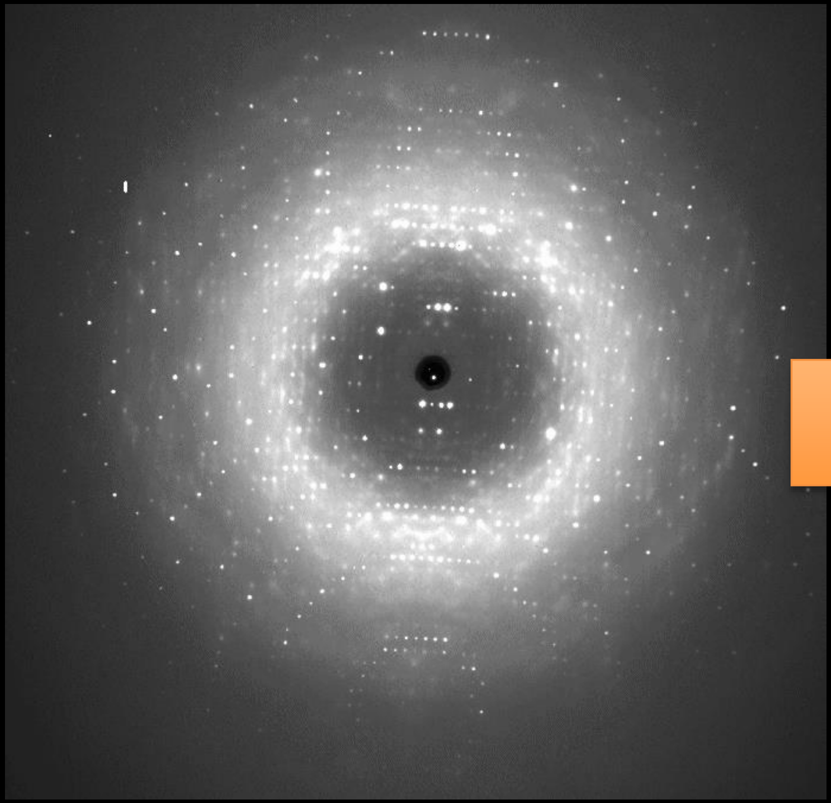
$$I = I_D + I_B$$

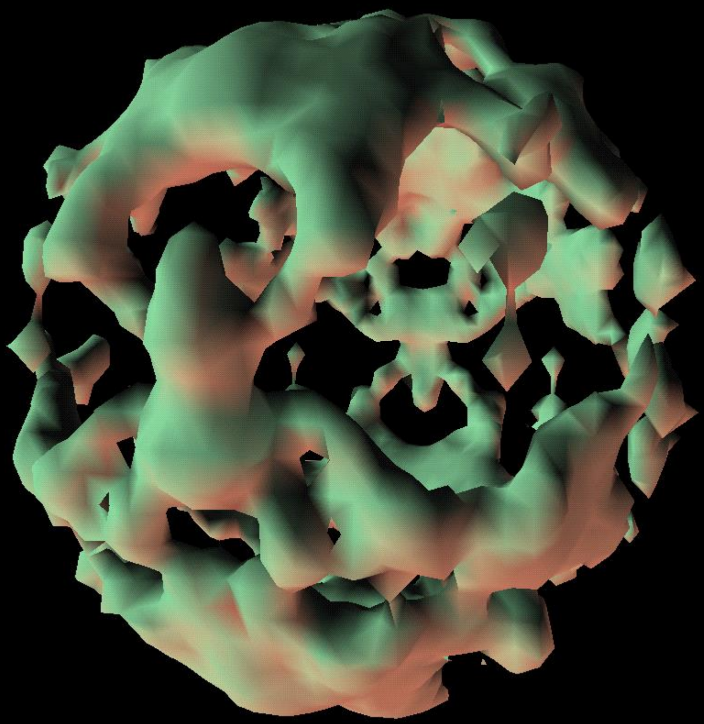
$$I_B = |\langle f \rangle|^2 \sum_n \sum_m e^{i\mathbf{q}\cdot(\mathbf{R}_n - \mathbf{R}_m)}$$

$$I_D(\mathbf{q}) = N\left(\left\langle \left| f_n - \langle f_n \rangle \right|^2 \right\rangle_n \right) \propto N \int d\mathbf{x} e^{i\mathbf{q}\cdot\mathbf{x}} \int d\mathbf{x}' \left\langle \Delta\rho_n(\mathbf{x}')\Delta\rho_n(\mathbf{x}' + \mathbf{x}) \right\rangle_n$$

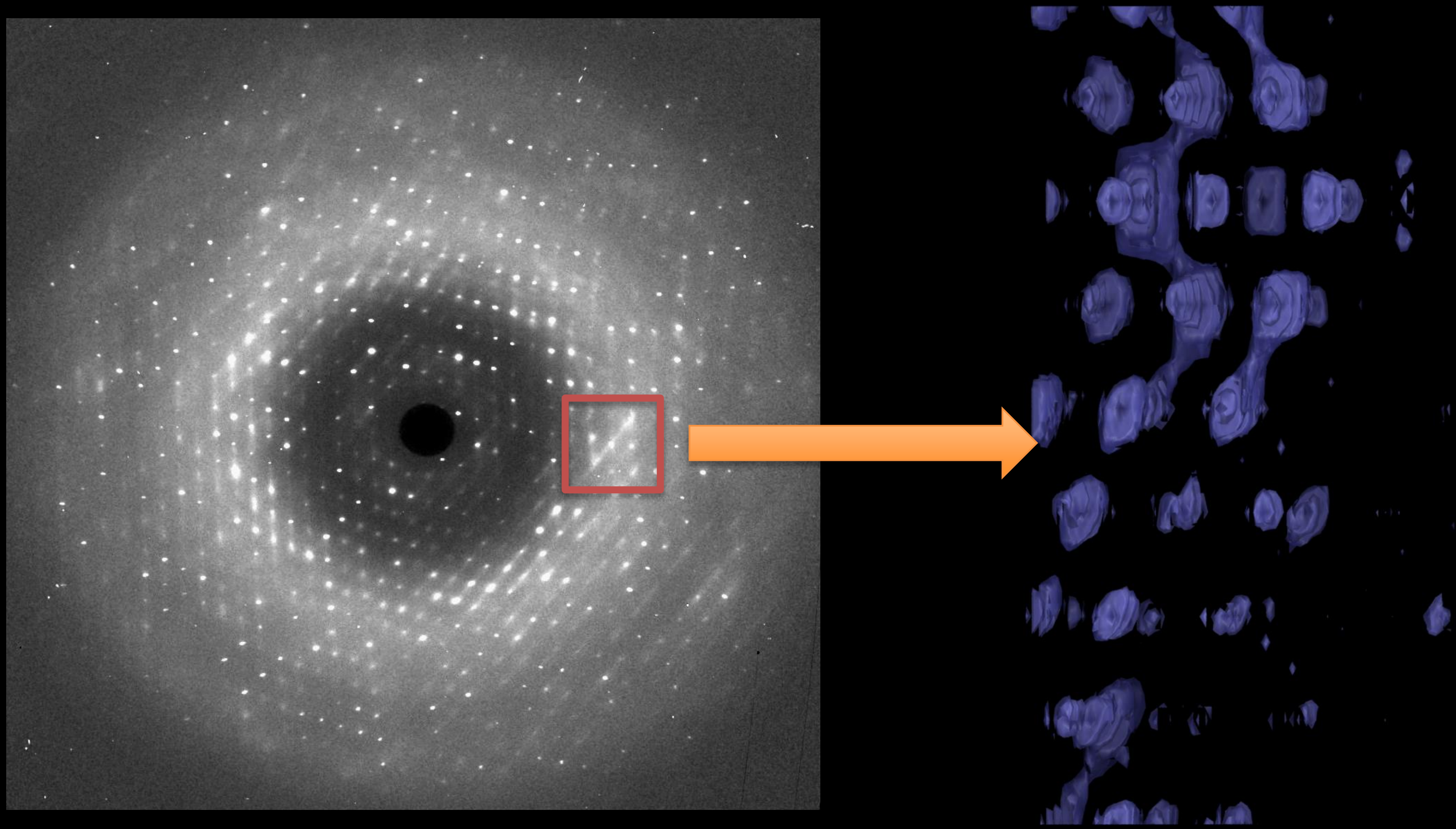
Sample using molecular dynamics simulations

Guinier 1956, 1963





Wall, 1996 (Ph.D. Thesis)
Wall, Ealick, and Gruner, *PNAS* 1997
Wall, *Methods Mol Biol* 2009
http://github.com/mewall/lunus



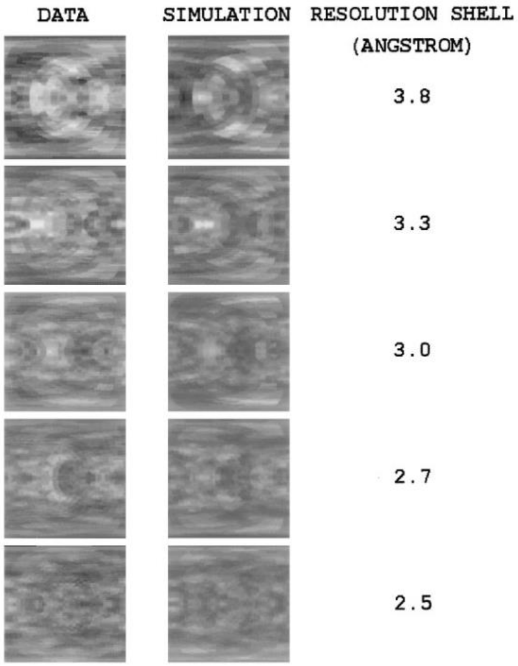
Wall, Clarage & Phillips, *Structure* 1997

$$I_D(\vec{q}) = (1 - e^{-\vec{q}\cdot\mathbf{U}\cdot\vec{q}})e^{-\vec{q}\cdot\mathbf{U}\cdot\vec{q}} I_0(\vec{q}) * \Gamma_{\mathbf{G}}(\vec{q})$$

$$\Gamma_{\mathbf{G}}(\vec{q}) = \frac{8\pi \det\|\mathbf{G}\|}{\left[1 + |\mathbf{G}\vec{q}|^2\right]^2}$$

Liquid-like motions
(Caspar et al)

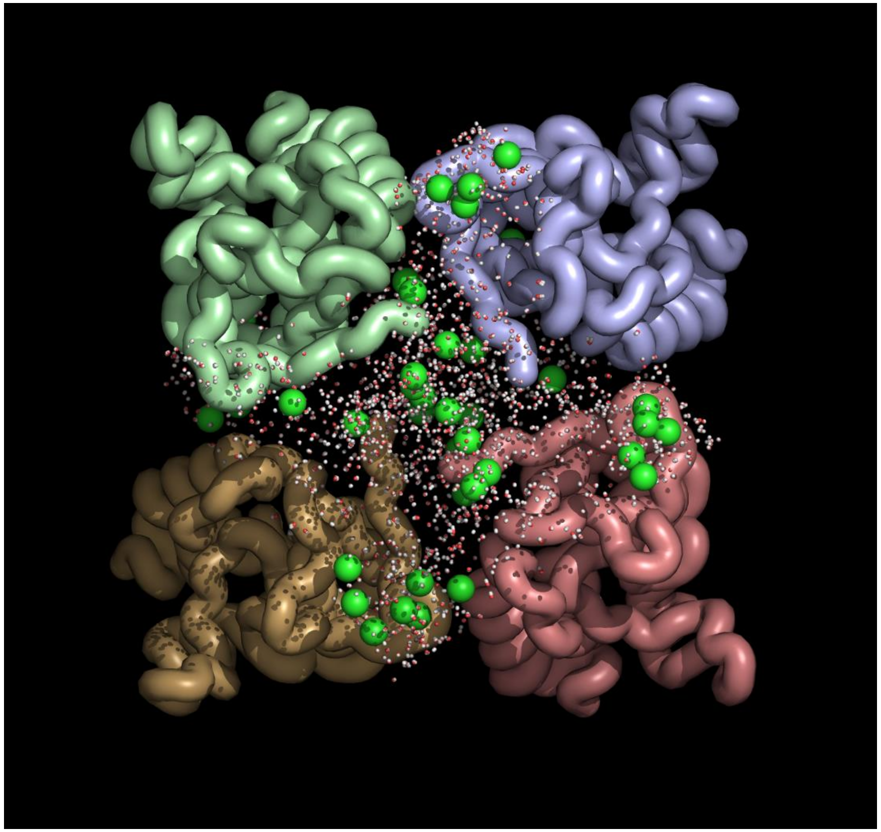
Snase: γ = 10 Å; σ = 0.36 Å
Calmodulin: γ = 4.8 Å; σ = 0.38 Å



$$\Gamma_{\mathbf{G}}(\vec{q}) = \frac{4\pi \det\|\mathbf{G}\|}{1 + |\mathbf{G}\vec{q}|^2}$$

Acoustic modes

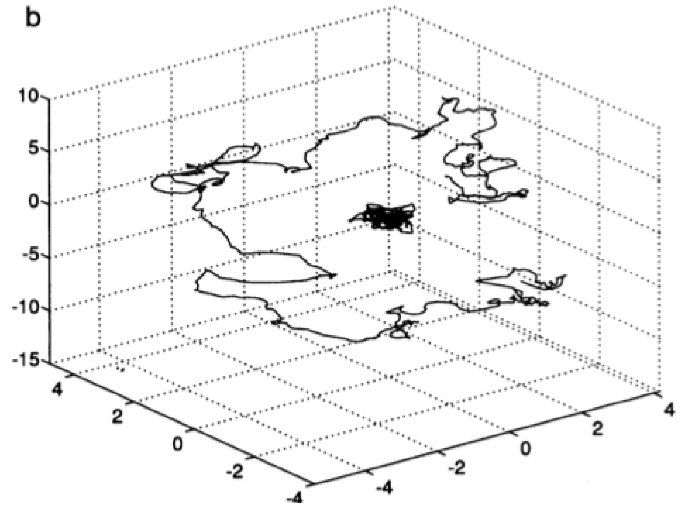
Correlations:	Displacements:
γ_1 = 50 Å	σ_1 = 0.0 Å
γ_2 = 135 Å	σ_2 = 0.4 Å
γ_3 = 85 Å	σ_3 = 0.0 Å



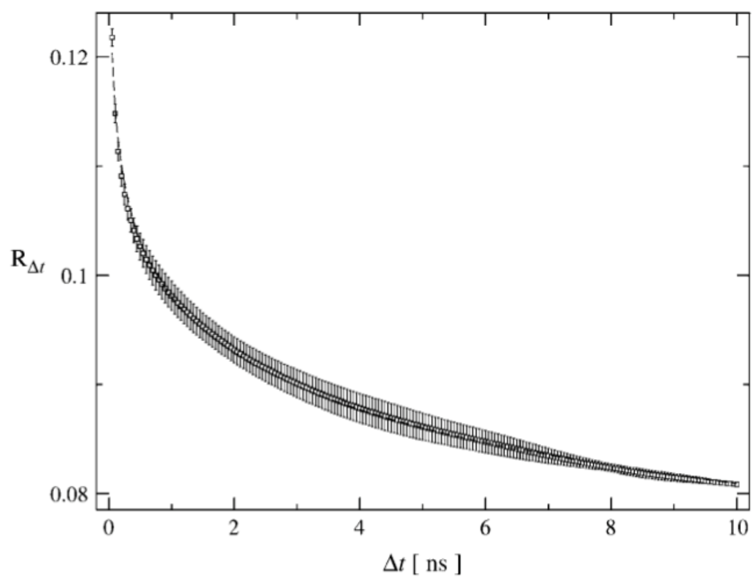
$$L = \frac{1}{2}\sum m_i \dot{\mathbf{R}}_i^2 - U\left(\mathbf{R}_1, \ldots, \mathbf{R}_N\right)$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{R}_k}\right) = -\frac{\partial L}{\partial R_k}$$

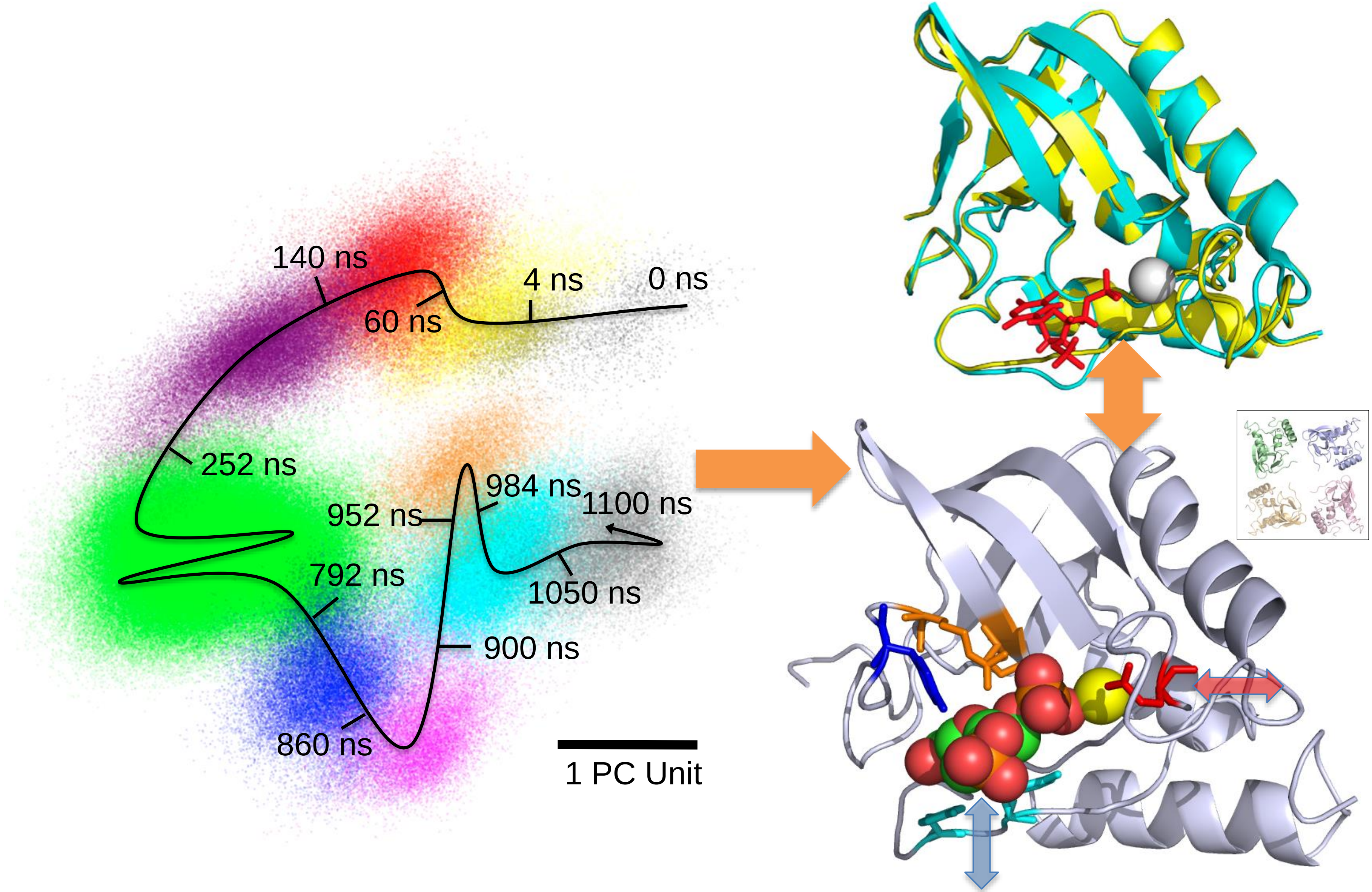
$$\mathbf{f}_i = m_i \ddot{\mathbf{R}}_i = -\frac{\partial U}{\partial \mathbf{R}_i}$$



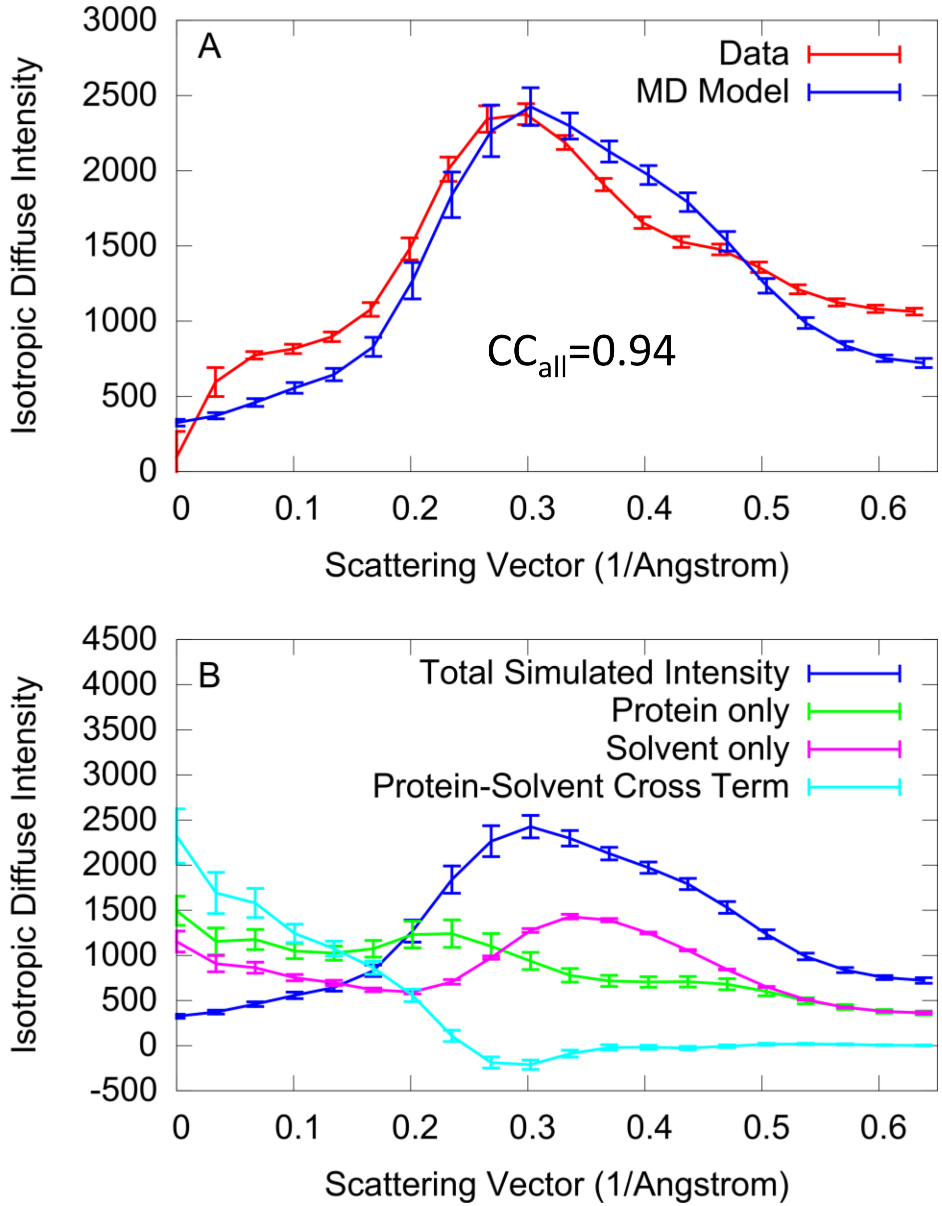
Clarage et al, *PNAS* 1995



Meinhold & Smith, *Biophys J* 2005

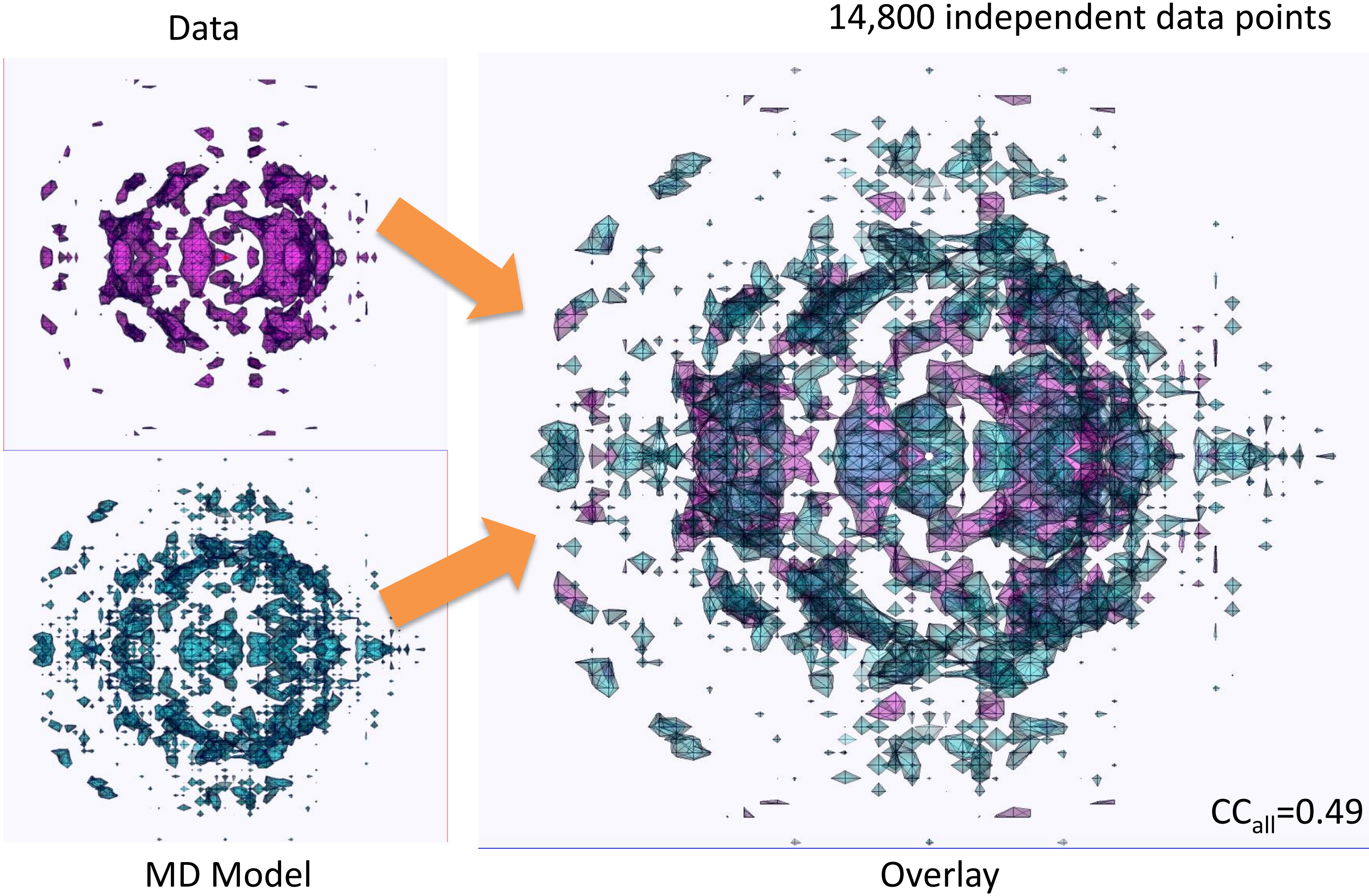


Wall et al, *PNAS* 2014



A
Data
MD Model
CC$_{all}$=0.94
Isotropic Diffuse Intensity
Scattering Vector (1/Angstrom)
B
Total Simulated Intensity
Protein only
Solvent only
Protein-Solvent Cross Term
Isotropic Diffuse Intensity
Scattering Vector (1/Angstrom)

14,800 independent data points

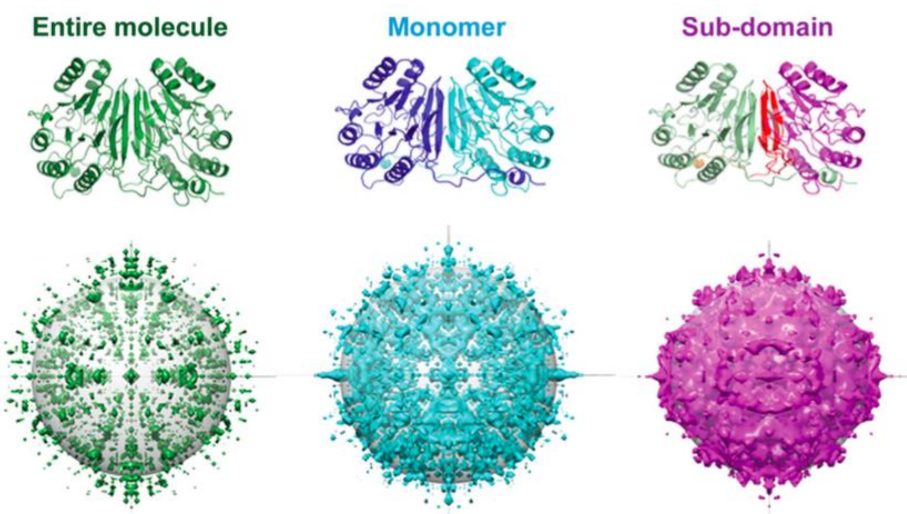


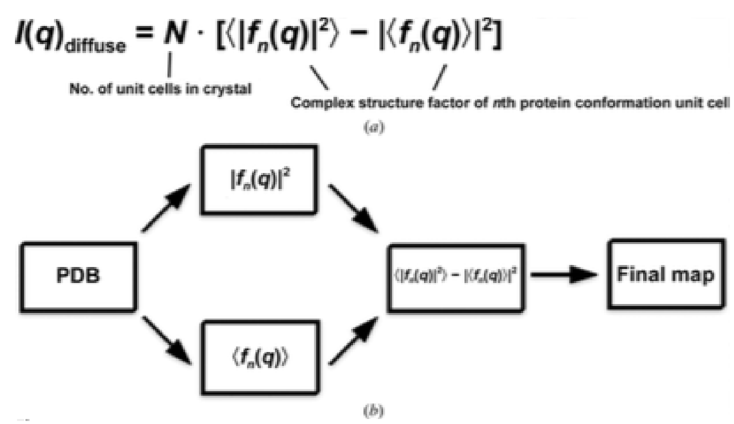
Overlay

Predicting X-ray diffuse scattering from translation–libration–screw structural ensembles

Andrew H. Van Benschoten,[a] Pavel V. Afonine,[b] Thomas C. Terwilliger,[c] Michael E. Wall,[d] Colin J. Jackson,[e] Nicholas K. Sauter,[b] Paul D. Adams,[b,f] Alexandre Urzhumtsev[g,h] and James S. Fraser[a*]

Acta Cryst. (2015). D**71**, 1657–1667





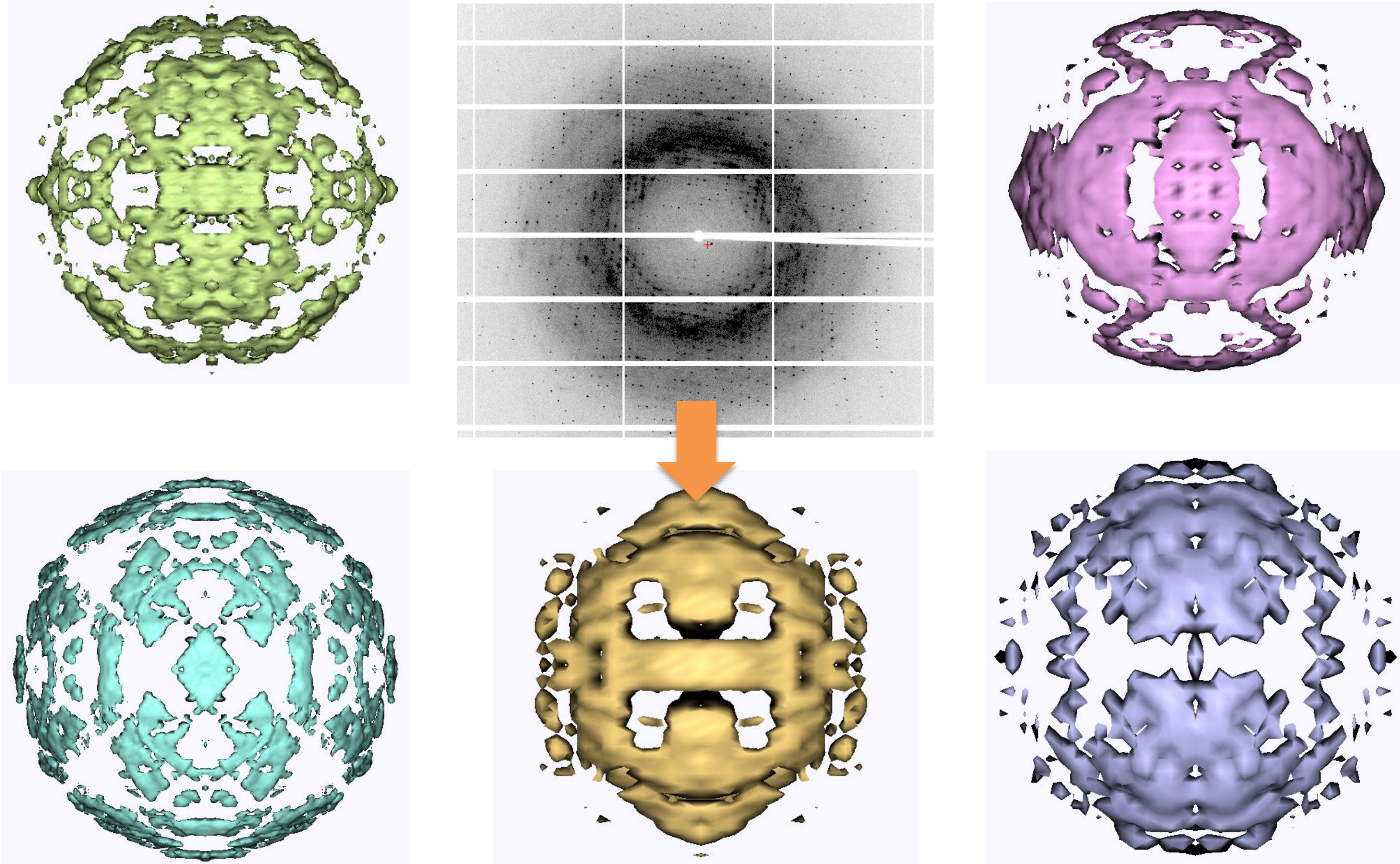


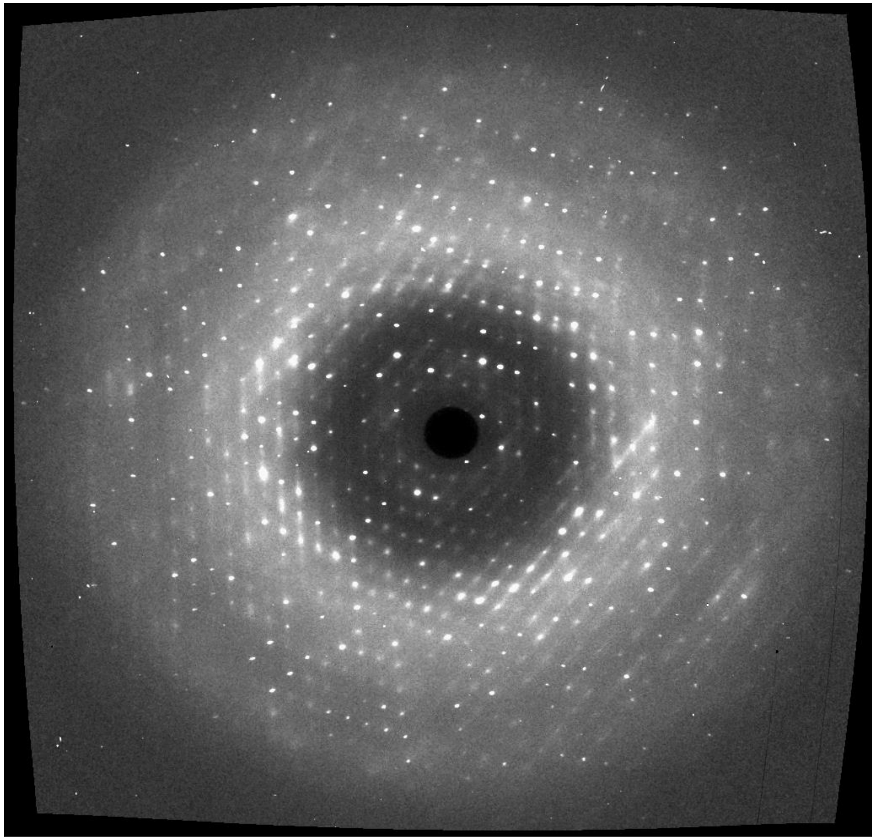
Image processing and scripts in Lunus (http://github.com/mewall/lunus)

DIALS indexing methods from Nicholas Sauter and Aaron Brewster, LBNL

Experiments of James Fraser and Others

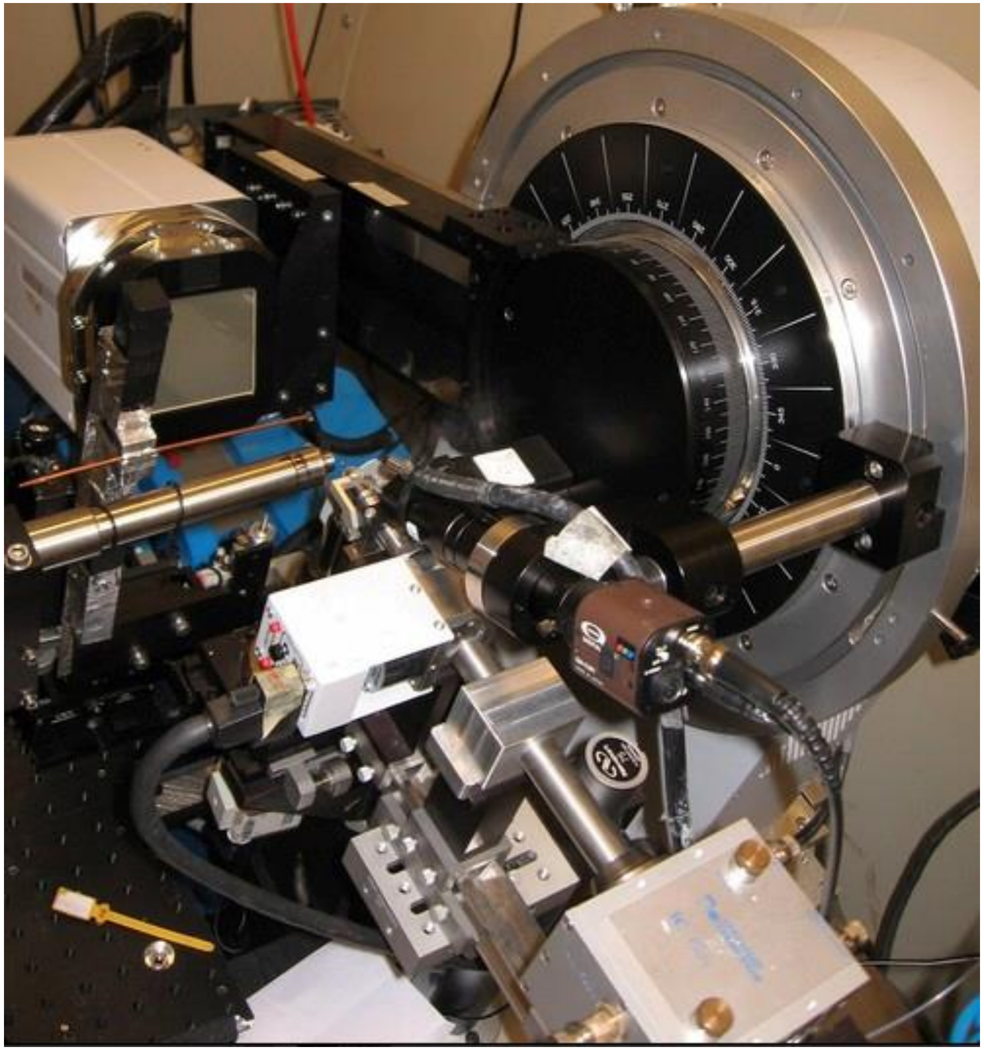
Raw Images

- No compression
- Simple layout of data
- Shared conventions for *r,c->x,y*
- Human-readable header
  - *i.e.* SMV or the like



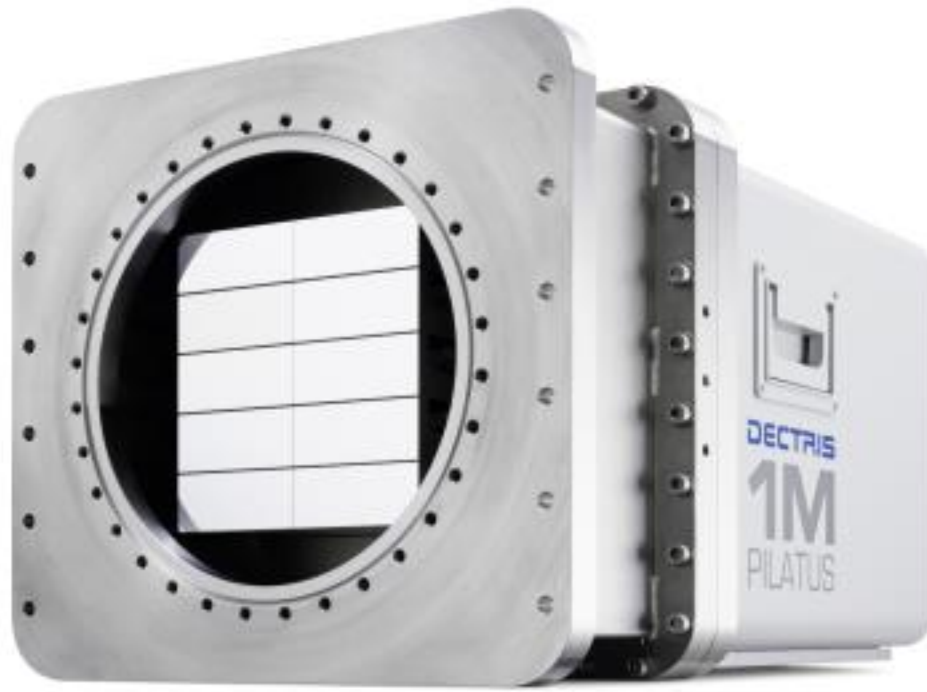
Beam Metadata

- Beam line
- Wavelength
  - Spectrum
- Polarization
  - Evidence
- Beam center
  - Evidence

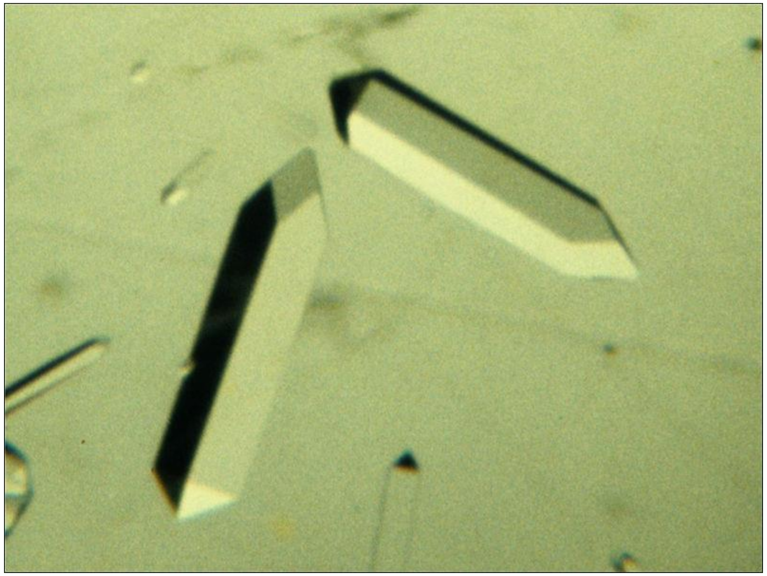


Detector Metadata

- Detector model
- Detailed operating mode
- Relation of ADU to X-ray counts
- Distance
  - Evidence
- Detector face rotation
  - Evidence
- Pixel size



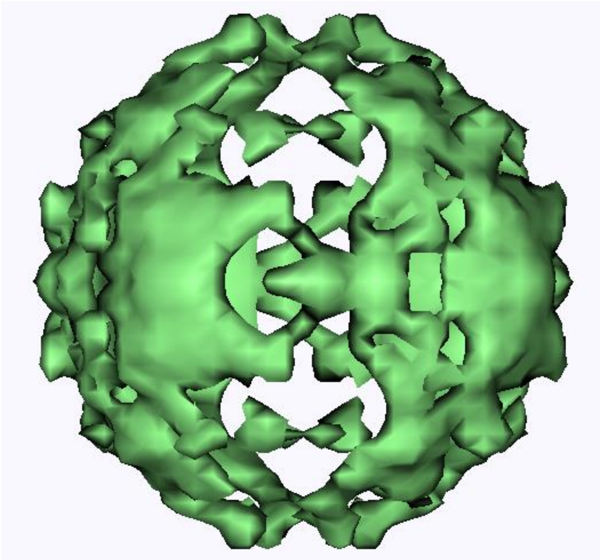
Crystal Metadata



- Space group
  - Evidence
- Unit cell
  - Evidence
- Chemical contents
- Light microscopy image for each exposure
  - Distinguish crystal from other scattering sources
  - Tomography model of specimen

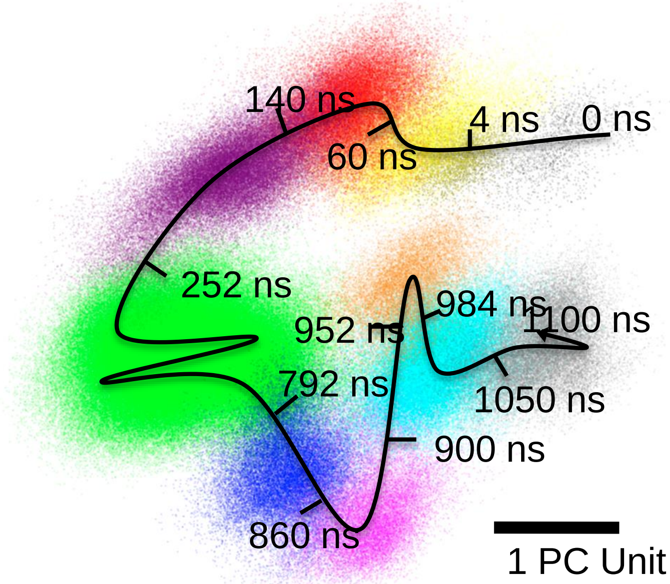
Integrated Diffuse Data Deposition

- *D*(*hkl*)
  - Fractional *hkl* possible
- Image processing parameters
  - Beam polarization
  - Solid-angle normalization
  - Bragg peak filtering
- Scale factors
- Frame-by-frame indexing information
- Flexible with respect to future needs for combined integration of Bragg and diffuse data

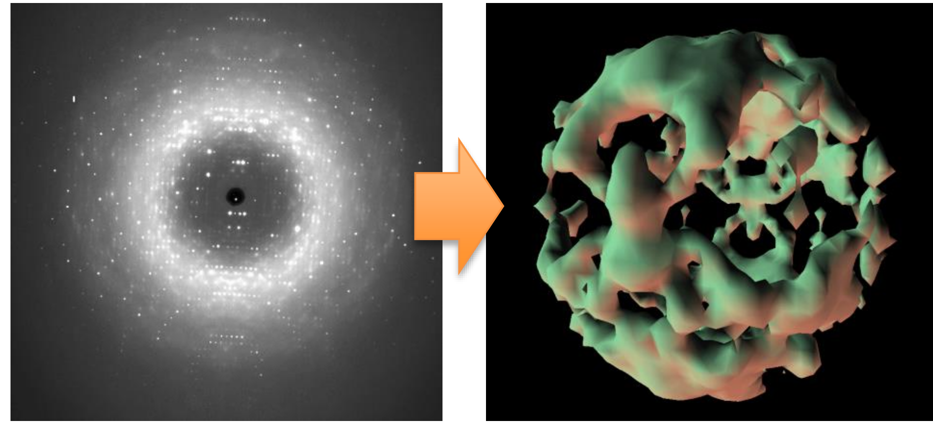


Diffuse Scattering Model Deposition

- Dynamical parameter values
  - Displacement correlations
  - Displacement amplitudes
  - Dispersion relation
- MD trajectories (large!)
- Calculated diffuse intensities



Acknowledgments



- Rice
  - George Phillips, Jr.
  - James Clarage (now at St Thomas)
- Cornell
  - Sol Gruner (formerly at Princeton)
  - Steven Ealick
  - CHESS staff
- FSU
  - Donald Caspar
- UCSF
  - James Fraser
  - Andrew Vanbenschoten
- LBNL
  - Paul Adams
  - Nicholas Sauter
  - Aaron Brewster
- LANL
  - Tom Terwilliger
- Global Phasing
  - Gérard Bricogne
- Funding
  - NSF, Welch Foundation, Keck Foundation, LANL LDRD
- ECM attendance
  - John Helliwell
  - Brian McMahon

Further Reading

- Amorós JL & Amorós M (1968) *Molecular Crystals; Their Transforms and Diffuse Scattering* (Wiley, New York).
- Caspar DL, Clarage J, Salunke DM, & Clarage M (1988) Liquid-like movements in crystalline insulin. *Nature* 332:659-662.
- Chacko S & Phillips GN, Jr. (1992) Diffuse X-ray scattering from tropomyosin crystals. *Biophys J* 61:1256-1266.
- Clarage JB, Clarage MS, Phillips WC, Sweet RM, & Caspar DL (1992) Correlations of atomic movements in lysozyme crystals. *Proteins* 12:145-157.
- Clarage JB, Romo T, Andrews BK, Pettitt BM, & Phillips GN, Jr. (1995) A sampling problem in molecular dynamics simulations of macromolecules. *Proc Natl Acad Sci U S A* 92:3288-3292.
- Doucet J & Benoit JP (1987) Molecular dynamics studied by analysis of the X-ray diffuse scattering from lysozyme crystals. *Nature* 325:643-646.
- Faure P*, et al.* (1994) Correlated intramolecular motions and diffuse X-ray scattering in lysozyme. *Nat Struct Biol* 1:124-128.
- Glover ID, Harris GW, Helliwell JR, & Moss DS (1991) The variety of X-ray diffuse-scattering from macromolecular crystals and its respective components. *Acta Crystallogr B* 47:960-968.
- Guinier A (1963) *X-ray Diffraction in Crystals, Imperfect Crystals, and Amorphous Bodies* (W. H. Freeman and Company, San Francisco).
- Helliwell JR, Glover ID, Jones A, Pantos E, & Moss DS (1986) Protein dynamics - use of computer-graphics and protein crystal diffuse-scattering recorded with synchrotron X-radiation. *Biochem Soc Transact* 14:653-655.
- Héry S, Genest D, & Smith JC (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* (Bell, London).
- Kolatkar AR, Clarage JB, & Phillips GN, Jr. (1994) Analysis of diffuse scattering from yeast initiator tRNA crystals. *Acta Crystallogr D* 50:210-218.
- Lonsdale K (1942) X-ray study of crystal dynamics : An historical and critical survey of experiment and theory. Proceedings of the Physical Society 54:314-353.
- Meinhold L & Smith JC (2005) Fluctuations and correlations in crystalline protein dynamics: a simulation analysis of Staphylococcal nuclease. *Biophys J* 88:2554-2563.
- Meinhold L & Smith JC (2005) Correlated dynamics determining X-ray diffuse scattering from a crystalline protein revealed by molecular dynamics simulation. *Phys Rev Lett* 95:218103.
- Meinhold L & Smith JC (2007) Protein dynamics from X-ray crystallography: anisotropic, global motion in diffuse scattering patterns. *Proteins* 66:941-953.
- Meinhold L, Merzel F, & Smith JC (2007) Lattice dynamics of a protein crystal. *Phys Rev Lett* 99:138101.
- Mizuguchi K, Kidera A, & Gō N (1994) Collective motions in proteins investigated by X-ray diffuse scattering. *Proteins* 18:34-48.
- Moore PB (2009) On the relationship between diffraction patterns and motions in macromolecular crystals. *Structure* 17:1307-1315.
- Phillips GN, Jr., Fillers JP, & Cohen C (1980) Motions of tropomyosin. Crystal as metaphor. *Biophys J* 32:485-502.
- Riccardi D, Cui Q, & Phillips GN, 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.
- Vanbenschoten AH et al (2015)Predicting X-ray diffuse scattering from translation– libration–screw structural ensembles. Acta Cryst D71: 1657.
- Wall ME, Adams PD, Fraser JS, & Sauter NK (2014) Diffuse X-ray scattering to model protein motions. *Structure* 22:182-184.
- Wall ME, Ealick SE, & Gruner SM (1997) Three-dimensional diffuse X-ray scattering from crystals of *Staphylococcal* nuclease. *Proc Natl Acad Sci U S A* 94:6180-6184.
- Wall ME, Clarage JB, & Phillips GN (1997) Motions of calmodulin characterized using both Bragg and diffuse X-ray scattering. Structure 5:1599-1612.
- Wall ME, et al. (2014) Conformational dynamics of a crystalline protein from microsecond-scale molecular dynamics simulations and diffuse X-ray scattering. Proc Natl Acad Sci U S A 111:17887- 17892.
- Warren BE (1969) *X-Ray Diffraction* (Addison-Wesley, Reading, MA).
- Welberry TR (2004) *Diffuse X-Ray Scattering and Models of Disorder* (Oxford University Press, Oxford).
- Willis BTM & Pryor AW (1975) *Thermal Vibrations in Crystallography* (Cambridge University Press, Cambridge).
- Wilson MA (2013) Visualizing networks of mobility in proteins. *Nat Meth* 10:835-837.
- Wooster WA (1962) *Diffuse X-Ray Reflections from Crystals* (Oxford University Press, Oxford).
- Zachariasen W (1945) *Theory of X-Ray Diffraction in Crystals* (Wiley, New York).