Dude, Where Are My X-Rays?
Real and Imagined Challenges in Diffuse Scattering

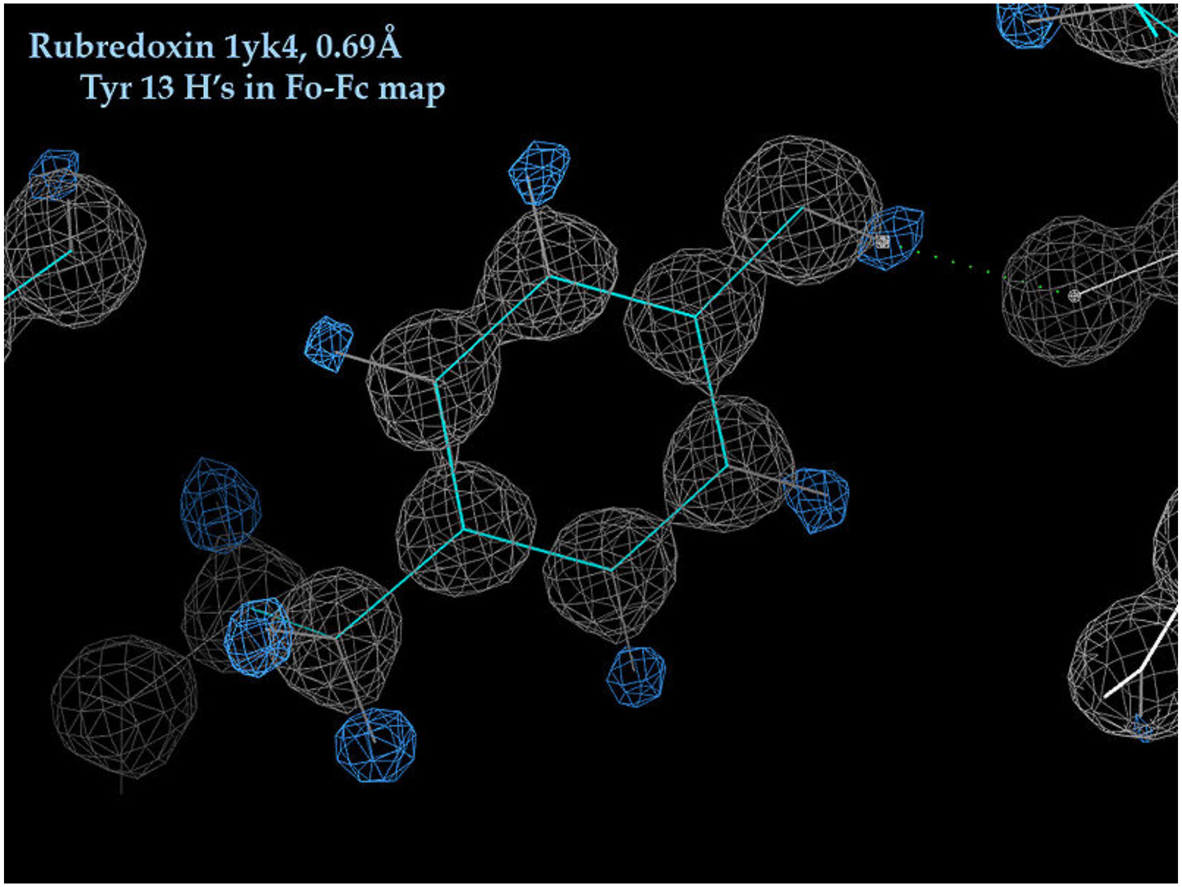
Michael Wall
Los Alamos National Laboratory

2013 ALS User Meeting
Diffuse Scattering Workshop
9 October 2013
LA-UR-13-27787

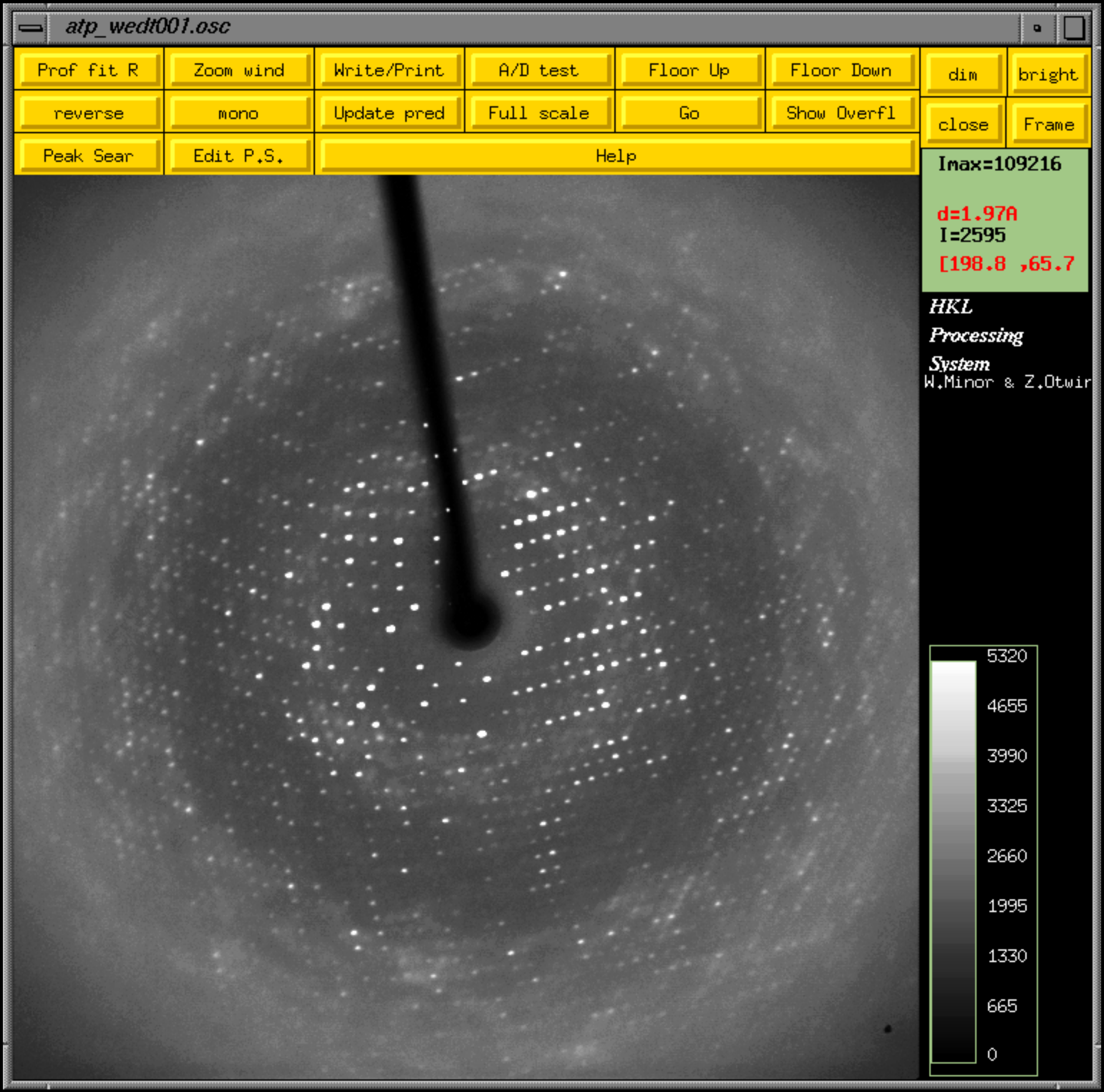
Protein Crystallography: What Information *Do* We Look For?

- High-Resolution structure
- Small B-factors

- Ideal: Perfect crystals.
- What is the reality?



http://commons.wikimedia.org/wiki/File:Protein_Hydrogens_in_electron_Density_1yk4_Y13.jpg#filelinks



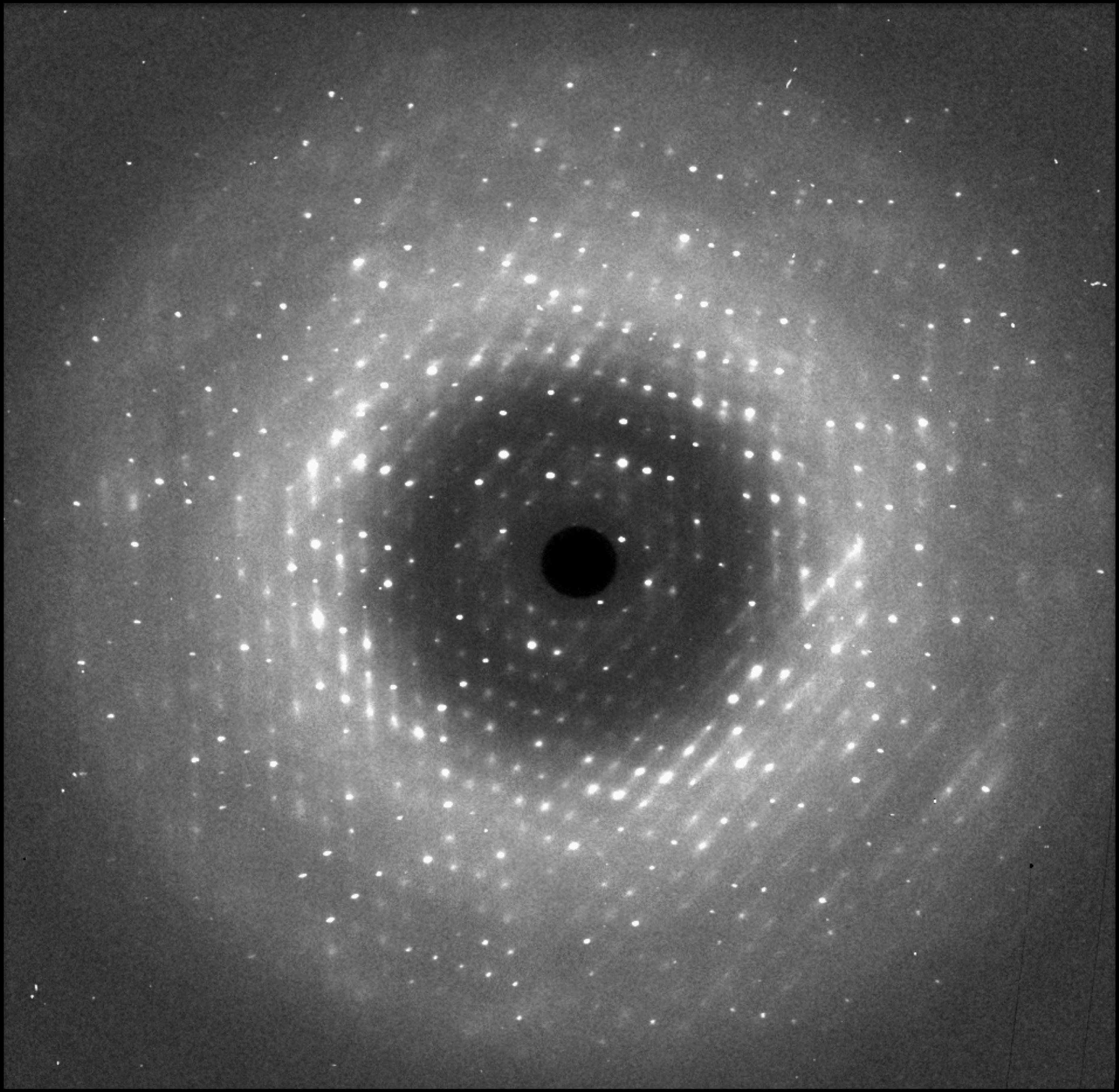
Imax=109216

d=1.97A
I=2595
[198.8 ,65.7

Image courtesy Mario Amzel

Protein Crystals Are Naturally Imperfect

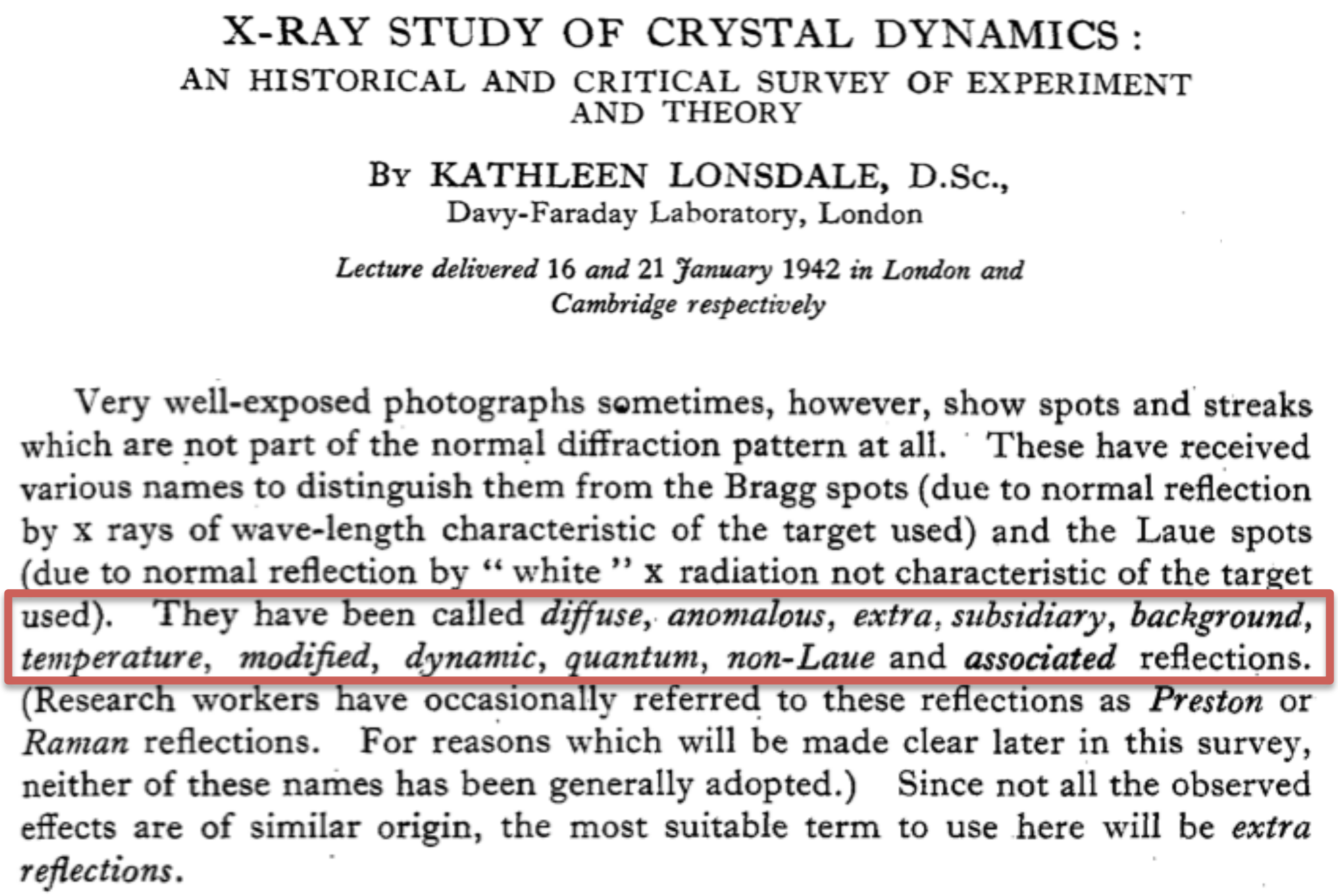
- Biologically relevant variations lead to imperfect crystals
- Current practice:
  - Seek perfect crystals, even if these variations are suppressed
    - Coax the missing X-rays back into the Bragg peaks
- Better practice:
  - Try to completely understand the crystal that Nature has given us
    - Use the "missing" X-rays right where they are – in the diffuse scattering



Interpretation of Diffuse Features

- *Size*
  - Large = correlations within unit cell
  - Small = correlations between unit cells
- *Streaking*
  - Direction of weak coupling/soft modes
- *Strength*
  - Relative intensity is highest along direction of motion

What Should We Call It?



X-RAY STUDY OF CRYSTAL DYNAMICS:

AN HISTORICAL AND CRITICAL SURVEY OF EXPERIMENT AND THEORY

By KATHLEEN LONSDALE, D.Sc.,
Davy-Faraday Laboratory, London

Lecture delivered 16 *and* 21 *January* 1942 *in London and Cambridge respectively*

Very well-exposed photographs sometimes, however, show spots and streaks which are not part of the normal diffraction pattern at all. These have received various names to distinguish them from the Bragg spots (due to normal reflection by x rays of wave-length characteristic of the target used) and the Laue spots (due to normal reflection by "white" x radiation not characteristic of the target used). They have been called *diffuse*, *anomalous*, *extra*, *subsidiary*, *background*, *temperature*, *modified*, *dynamic*, *quantum*, *non-Laue* and ***associated*** reflections. (Research workers have occasionally referred to these reflections as *Preston* or *Raman* reflections. For reasons which will be made clear later in this survey, neither of these names has been generally adopted.) Since not all the observed effects are of similar origin, the most suitable term to use here will be *extra reflections*.

Lonsdale, Proc Phys Soc 1942

THE CRYSTALLINE STATE—VOL II
Editor : SIR LAWRENCE BRAGG

THE OPTICAL PRINCIPLES
OF
THE DIFFRACTION OF X-RAYS

By

R. W. JAMES
D.SC., F.R.S.
EMERITUS PROFESSOR OF PHYSICS IN THE UNIVERSITY OF CAPE TOWN

3. THE EXPERIMENTAL STUDY OF DIFFUSE SCATTERING

Why isn't everyone already using diffuse scattering?

- The central theory was established decades ago
  - See, e.g., James (1948)
- It's already present in diffraction images
- It contains new information about proteins
  - Correlated motions
- However, there have been challenges
  - Real
  - Imagined

Diffuse Scattering Can’t Be Observed

- “According to diffraction theory the signal is too small”

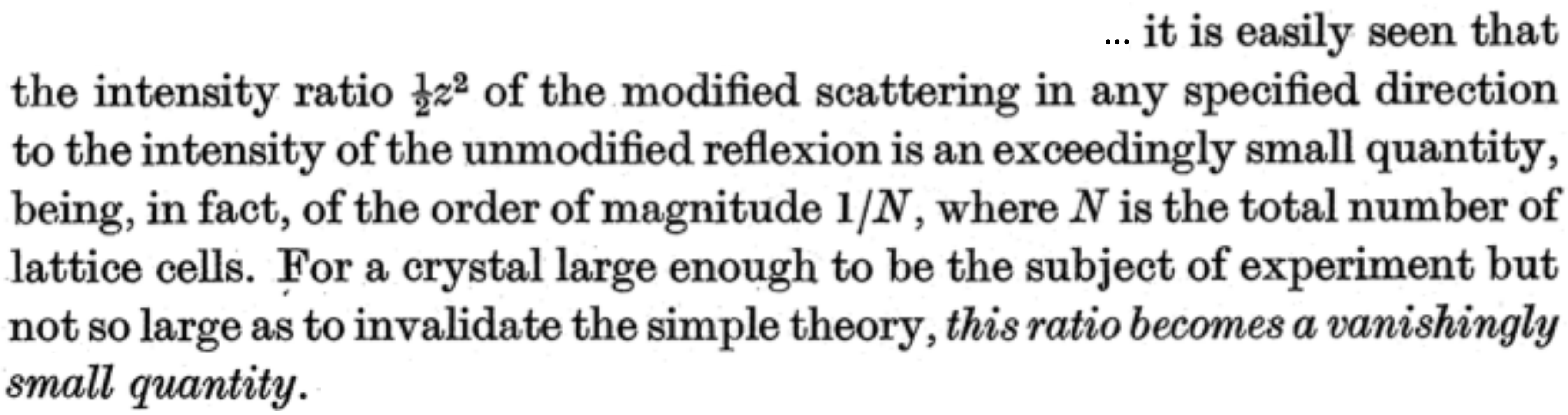
Diffuse Scattering Can't Be Observed

$$I_D = N\left(\left\langle \left|f\right|^2 \right\rangle - \left|\left\langle f \right\rangle\right|^2\right) \propto N$$

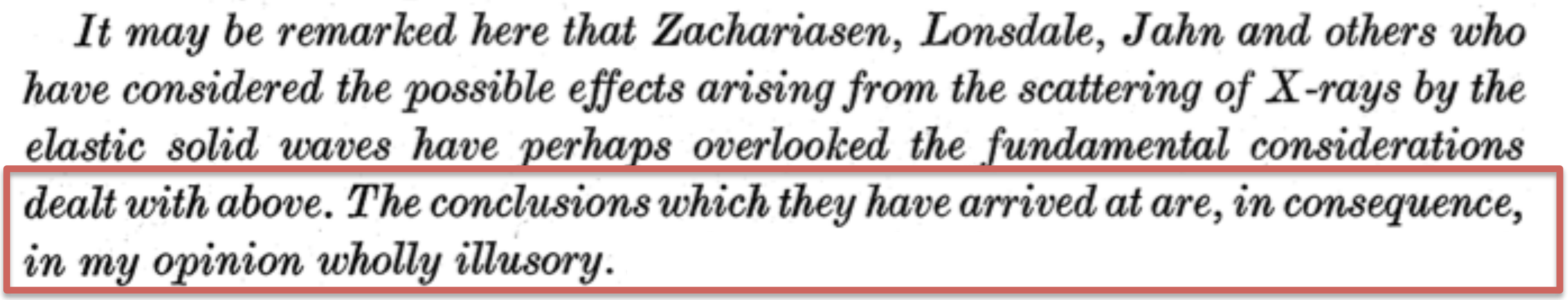
$$I_B = \left|\left\langle f \right\rangle\right|^2 \sum_n \sum_m e^{i\mathbf{q}\cdot\left(\mathbf{R}_n - \mathbf{R}_m\right)} \propto \prod_i \frac{\sin^2 N_i a_i s_i}{\sin^2 a_i s_i} \sim N^2$$

$$\Rightarrow I_B / I_D \sim N$$

Diffuse Scattering Can’t Be Observed



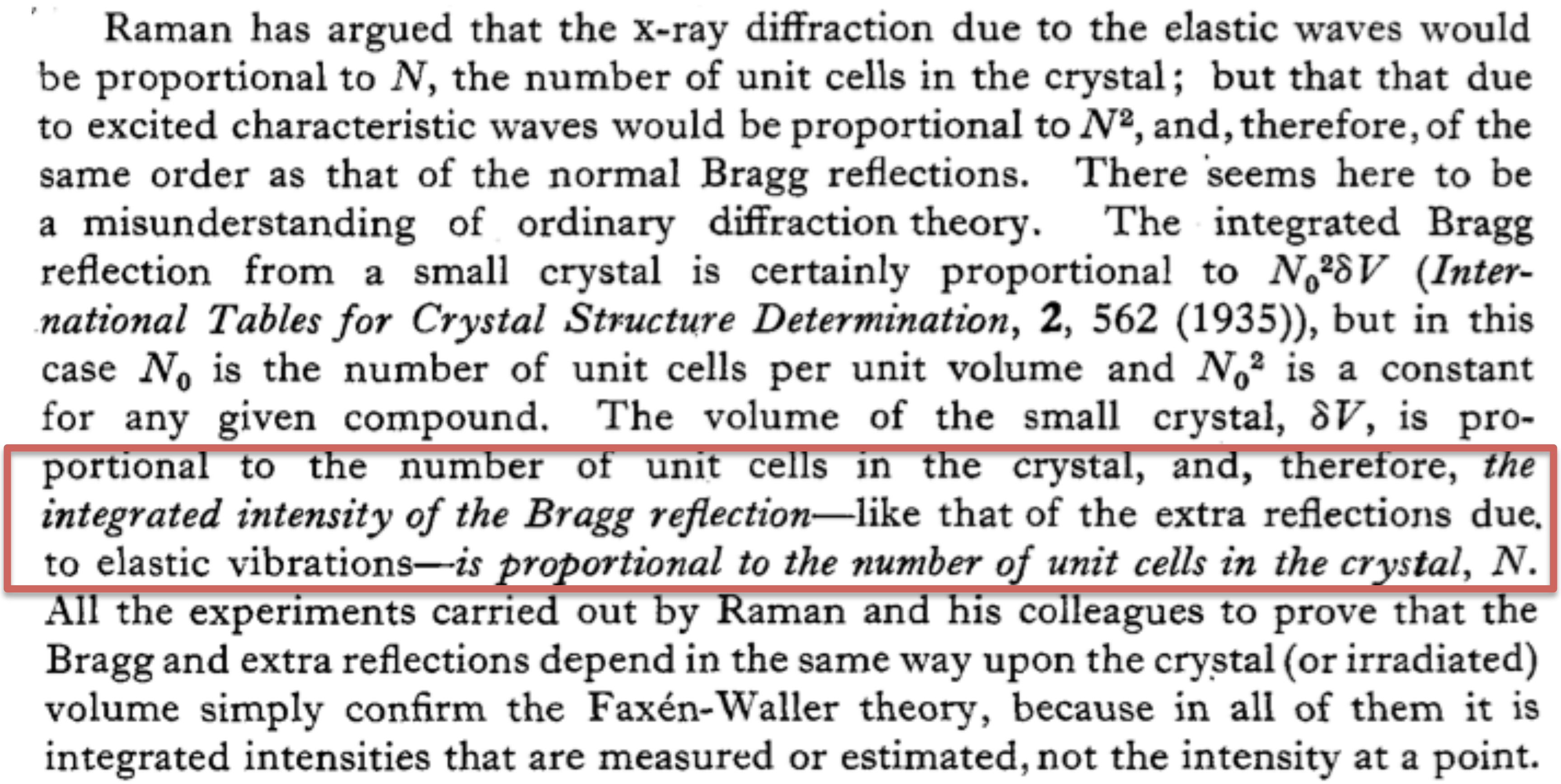
… it is easily seen that the intensity ratio $\frac{1}{2}z^2$ of the modified scattering in any specified direction to the intensity of the unmodified reflexion is an exceedingly small quantity, being, in fact, of the order of magnitude $1/N$, where N is the total number of lattice cells. For a crystal large enough to be the subject of experiment but not so large as to invalidate the simple theory, *this ratio becomes a vanishingly small quantity*.



It may be remarked here that Zachariasen, Lonsdale, Jahn and others who have considered the possible effects arising from the scattering of X-rays by the elastic solid waves have perhaps overlooked the fundamental considerations dealt with above. The conclusions which they have arrived at are, in consequence, in my opinion wholly illusory.

Raman, Proc Roy Soc Lond A 179 (1942) 289

Diffuse Scattering Can’t Be Observed



Raman has argued that the x-ray diffraction due to the elastic waves would be proportional to N, the number of unit cells in the crystal; but that that due to excited characteristic waves would be proportional to N^2, and, therefore, of the same order as that of the normal Bragg reflections. There seems here to be a misunderstanding of ordinary diffraction theory. The integrated Bragg reflection from a small crystal is certainly proportional to $N_0^2 \delta V$ (*International Tables for Crystal Structure Determination*, **2**, 562 (1935)), but in this case N_0 is the number of unit cells per unit volume and N_0^2 is a constant for any given compound. The volume of the small crystal, δV, is proportional to the number of unit cells in the crystal, and, therefore, *the integrated intensity of the Bragg reflection*—like that of the extra reflections due to elastic vibrations—*is proportional to the number of unit cells in the crystal*, N. All the experiments carried out by Raman and his colleagues to prove that the Bragg and extra reflections depend in the same way upon the crystal (or irradiated) volume simply confirm the Faxén-Waller theory, because in all of them it is integrated intensities that are measured or estimated, not the intensity at a point.

Lonsdale, Proc Phys Soc 54 (1942) 314

Diffuse Scattering ~~Can't~~ Can be Observed

$I_D \propto N$

I_B peak height $\sim N^2$

I_B peak width $\sim N_i^{-1}$

$\Rightarrow$ integrated peak intensity $I_B \sim N$

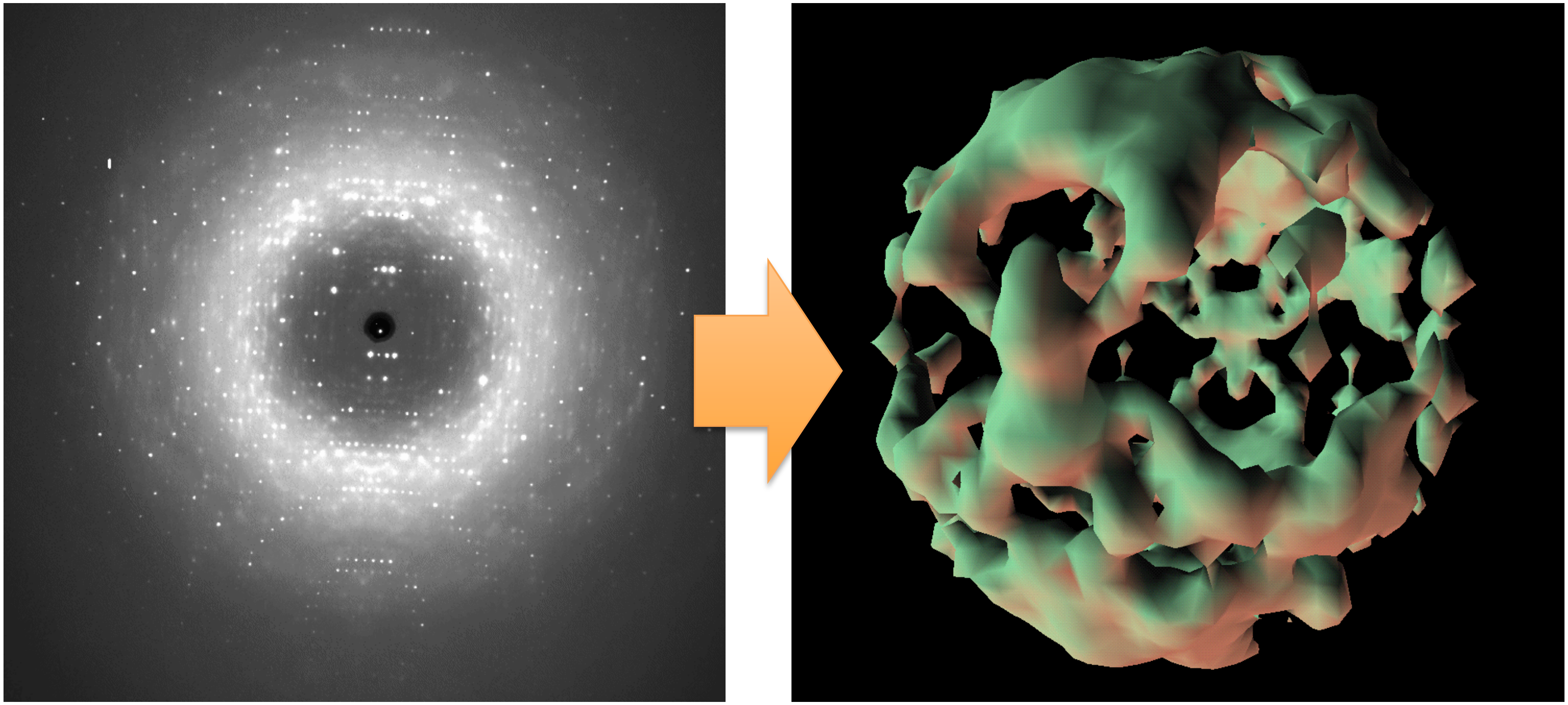
$$\Rightarrow I_B / I_D \sim 1$$

Imagined

Diffuse Scattering Can’t Be Measured

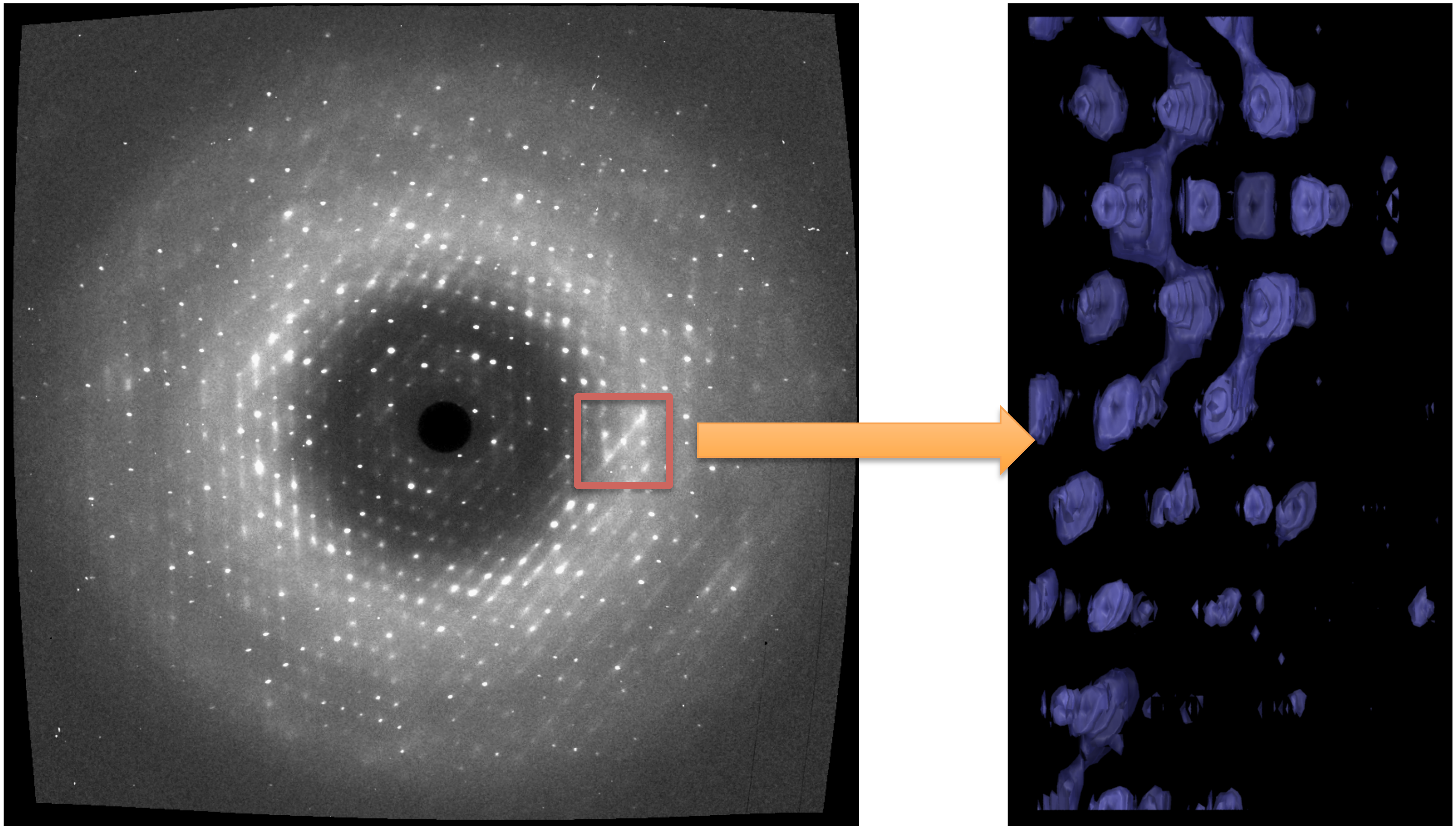
- “It can’t be observed”
- “Detectors aren’t good enough”

Diffuse Scattering from *Staph.* Nuclease Crystals



Wall, Ealick, and Gruner, PNAS 1997

Diffuse Scattering from Calmodulin



Wall, Clarage, and Phillips. Structure 1997

http://lunus.sourceforge.net

LUNUS

mewall

1 Download (This Week)

Last Update: 20 hours ago

Description

LUNUS software for generating 3D maps of diffuse scattering from macromolecular X-ray crystallography diffraction images. Also see http://lunus.sourceforge.net

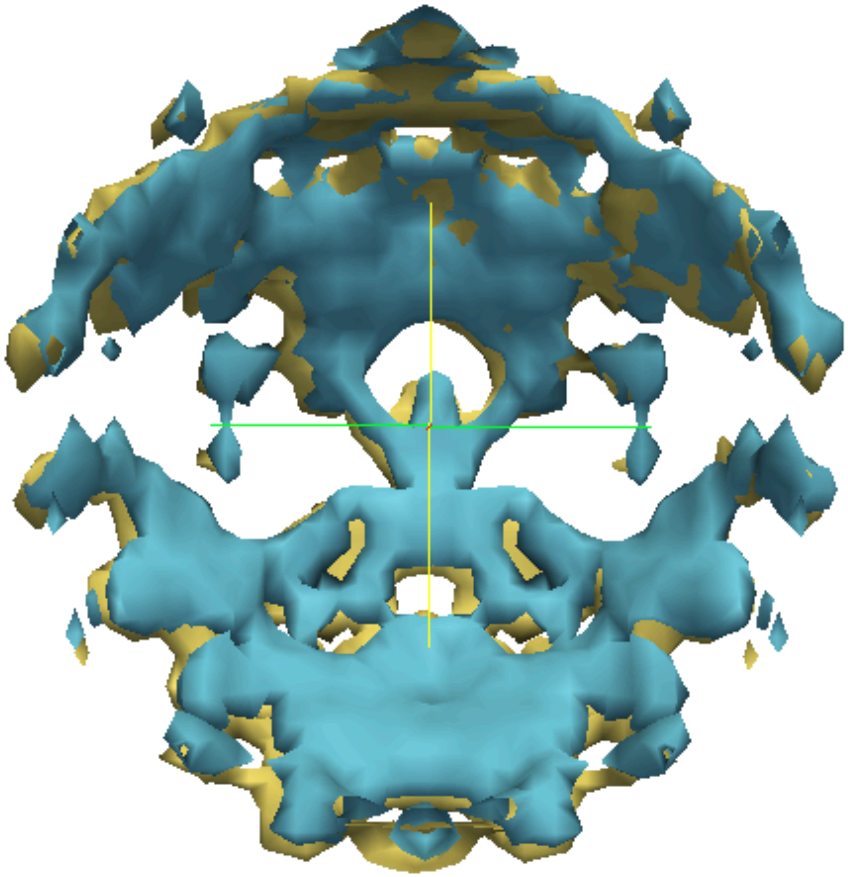
LUNUS Web Site ›

Categories: Physics

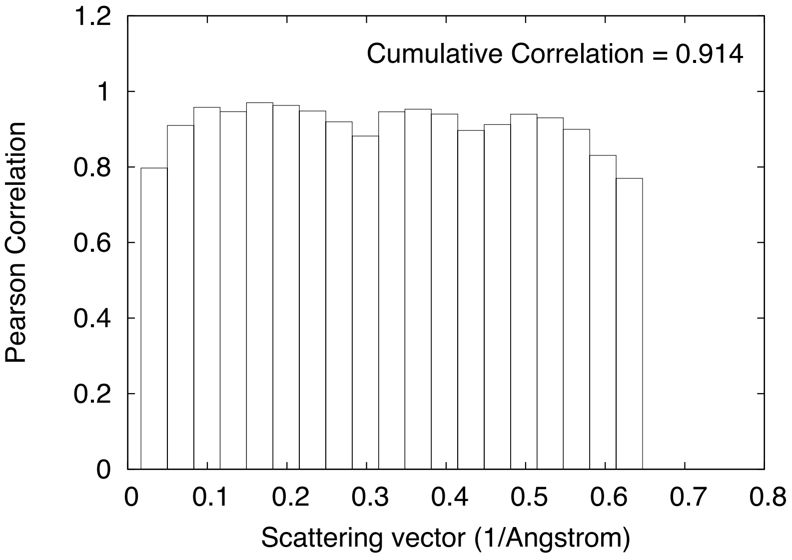
License: BSD License

Wall, Methods Mol Biol 2009

Lunus + CCTBX



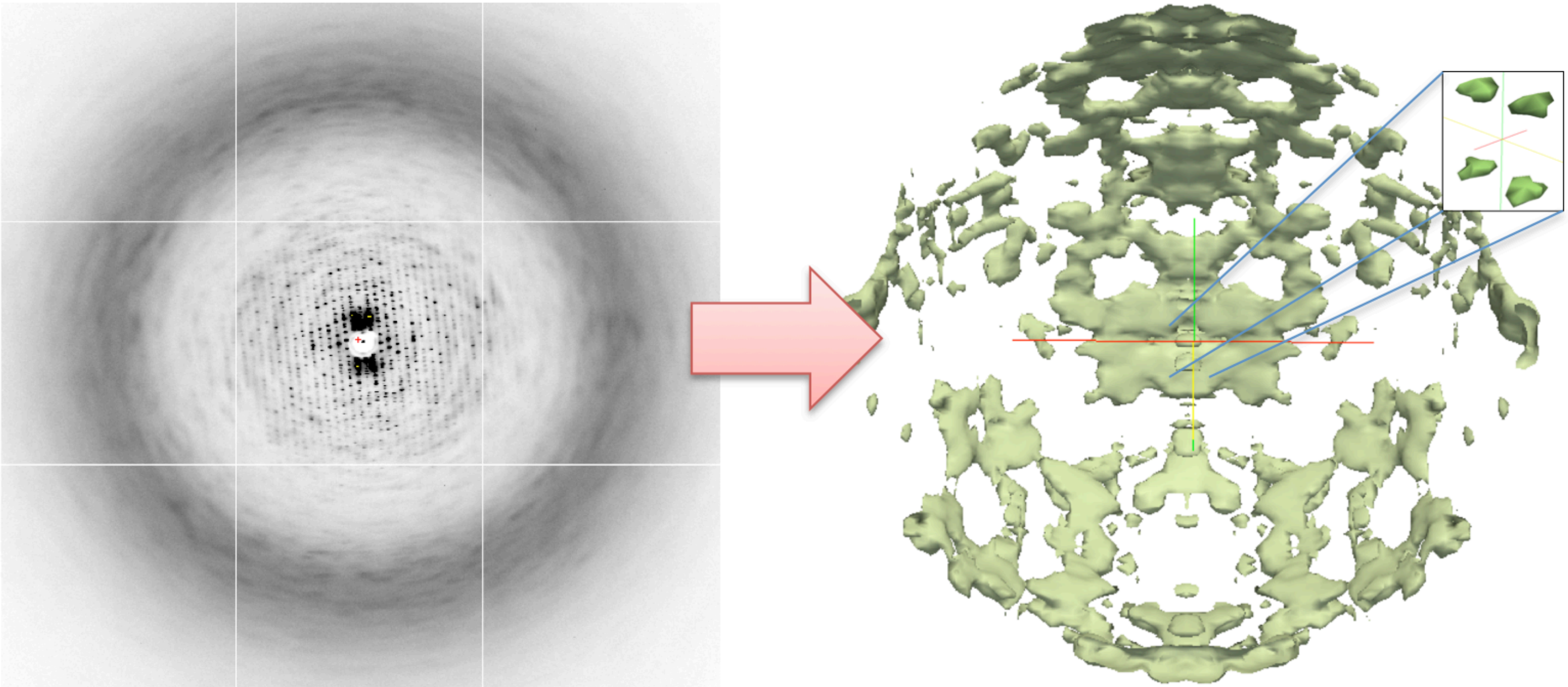
Correlation of all intensity = 0.998



Anisotropic Correlations

Using cctbx (Nicholas Sauter, LBNL)

Diffuse Scattering from Photosystem II



Data from Johan Hattne, Sauter Lab, LBNL

Using Lunus + cctbx

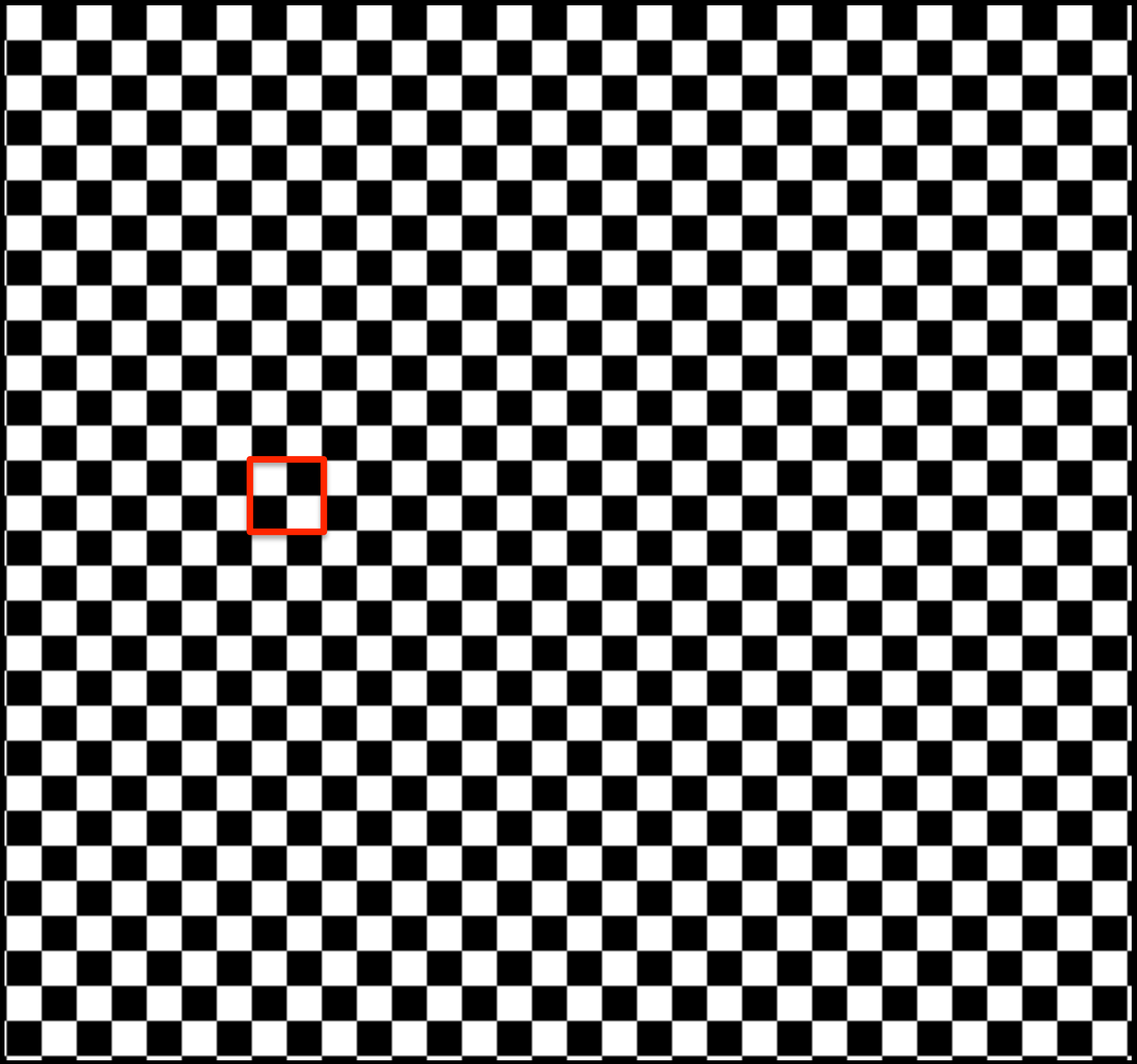
Diffuse Scattering ~~Can't~~ Can be Measured

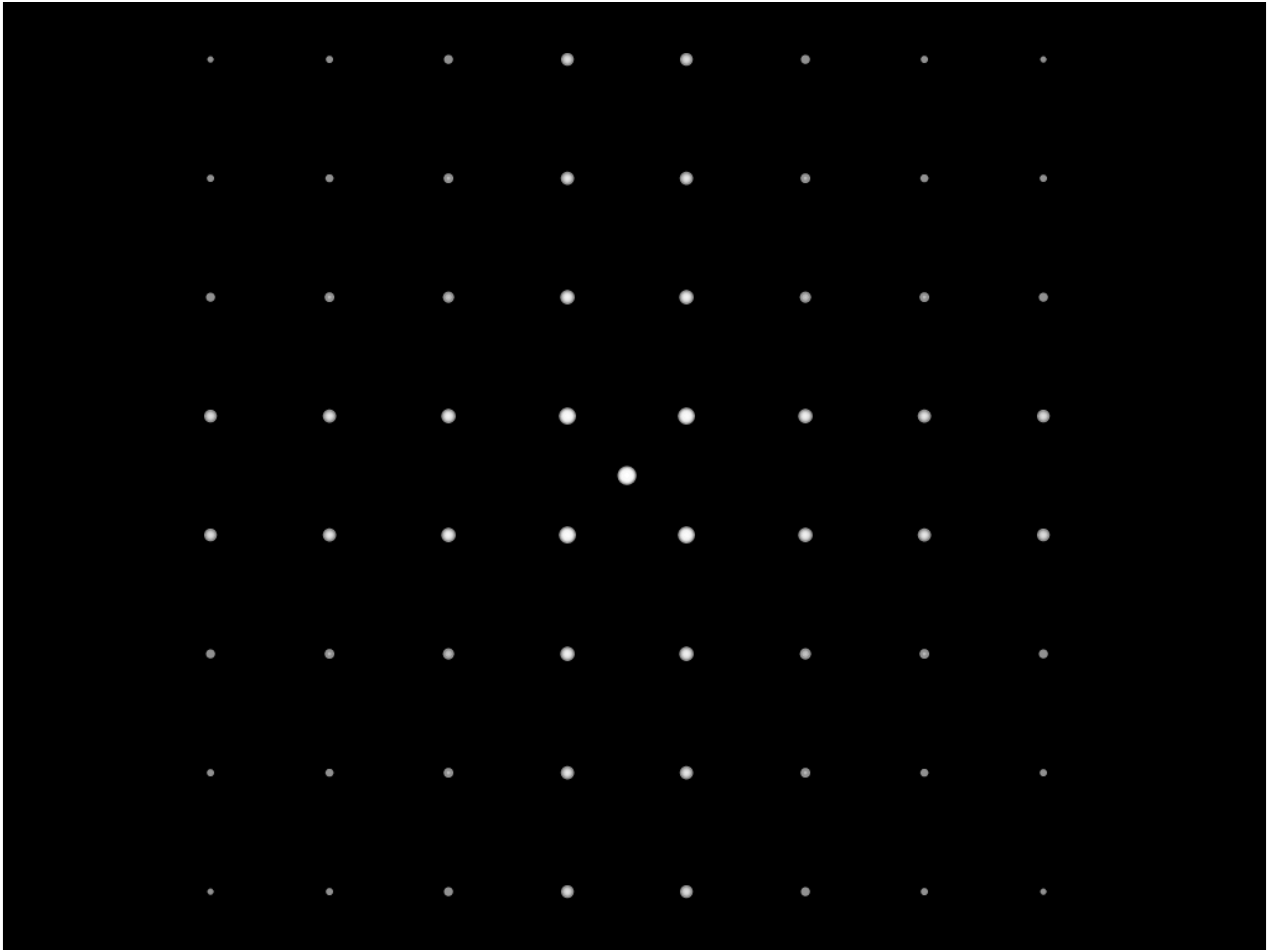
- 3D diffuse data have been measured for multiple systems
- These data are represented in the same way as the Bragg data
- Lunus software is publicly available

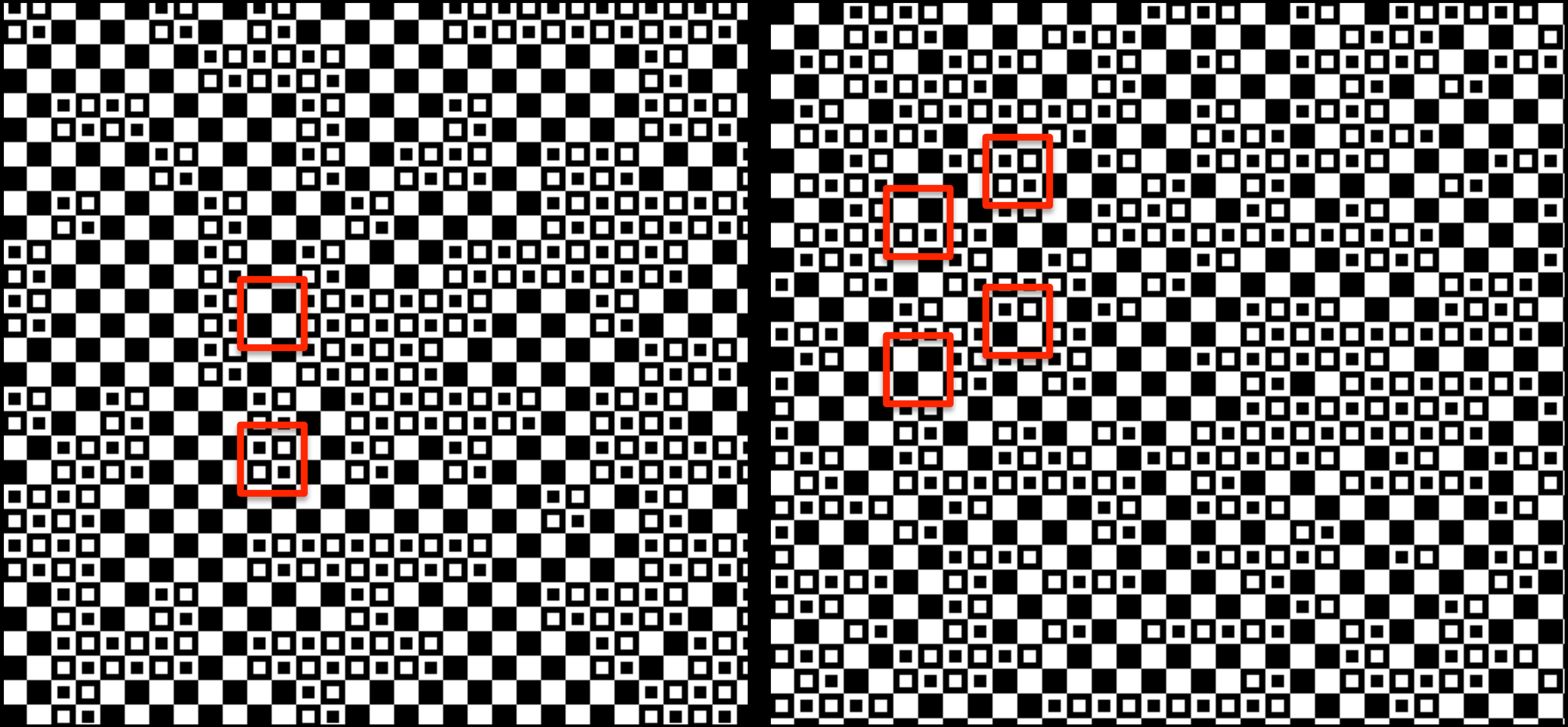
Imagined

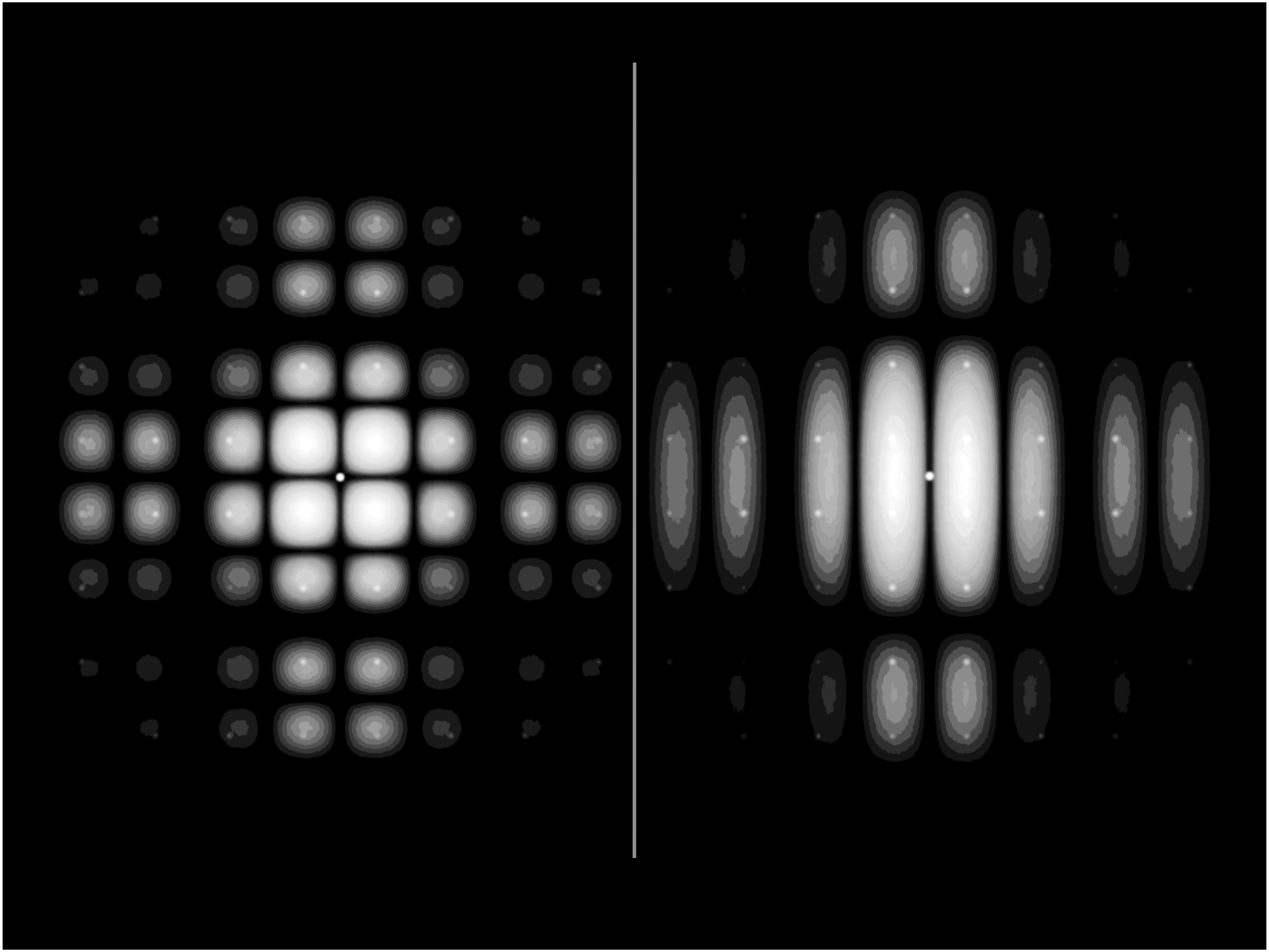
Diffuse Scattering Can’t Provide New Information

- “There’s nothing there that isn’t already in the Bragg peaks”









Diffuse Scattering ~~Can't~~ Can Provide New Information

- Information about correlated variations is in the diffuse scattering

$$I_D = N\left(\left\langle |f|^2 \right\rangle - \left|\left\langle f \right\rangle\right|^2\right)$$

- This information is not in the Bragg peaks

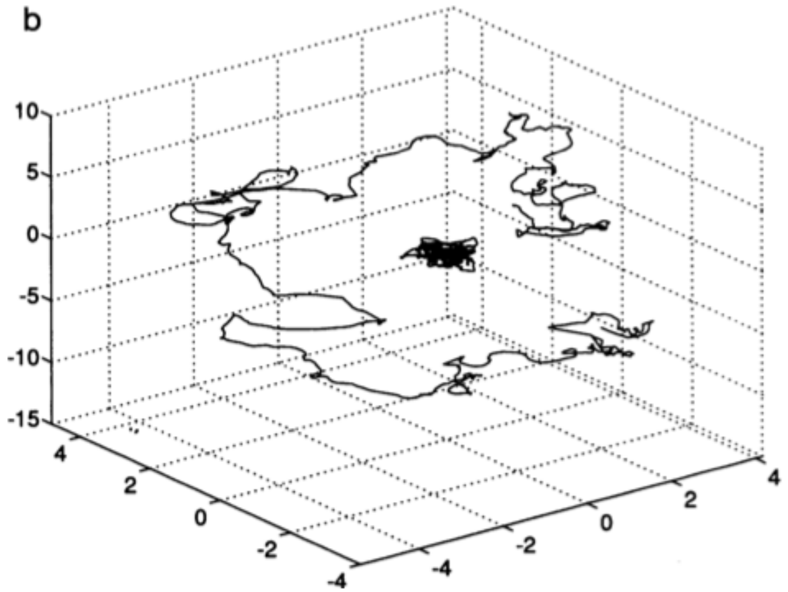
Imagined

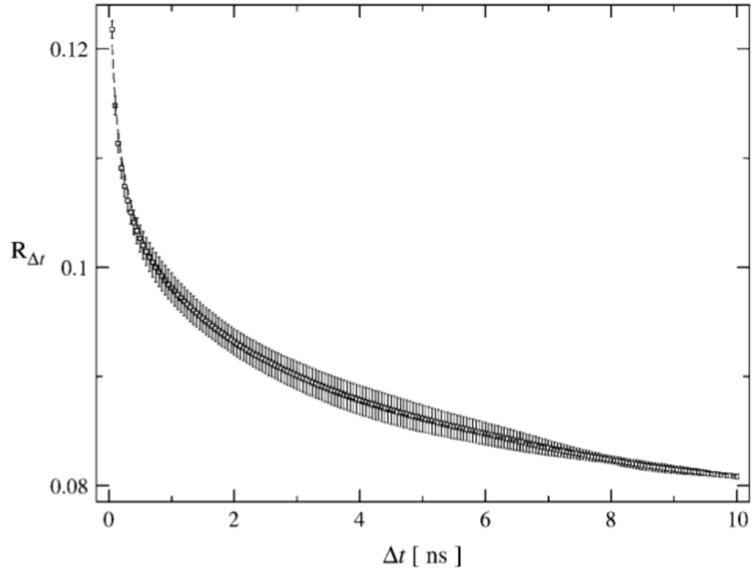
Diffuse Scattering Can’t Reveal Protein Dynamics

- Sampling problem in MD simulations
  - Can’t calculate two-point correlations
- Computationally intractable
  - Can’t calculate FT of a whole MD trajectory
- The model assumptions are wrong

Molecular Dynamics Sampling

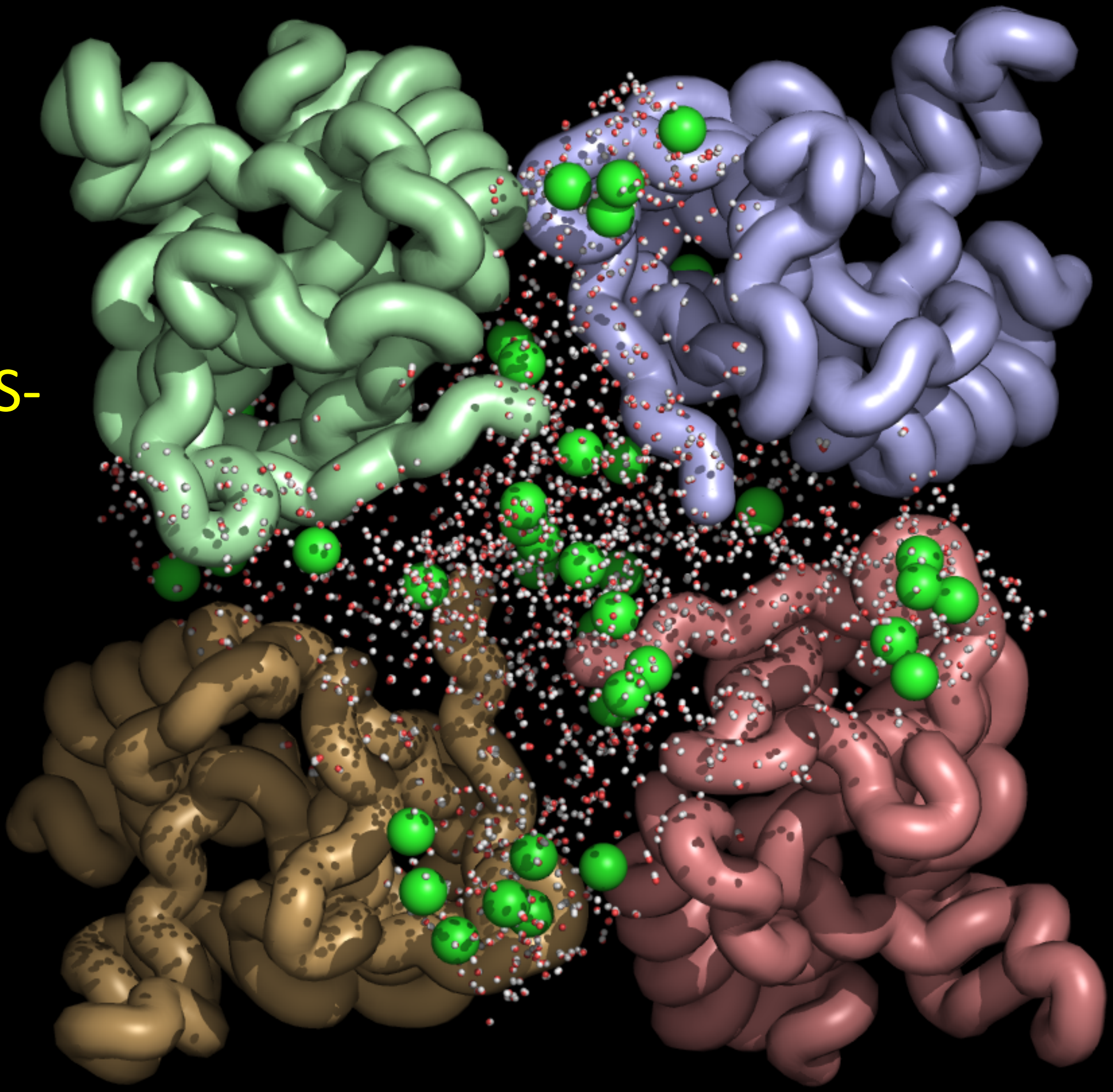
- Clarage et al., PNAS 1995
  - Myoglobin
  - 500 ps not long enough
  - Maybe 100 ns is?
- Meinhold and Smith, Biophys J 2005
  - *Staph*. nuclease
  - Improvement from 500 ps to 10 ns
  - Still not long enough
  - Maybe 1 microsecond is?

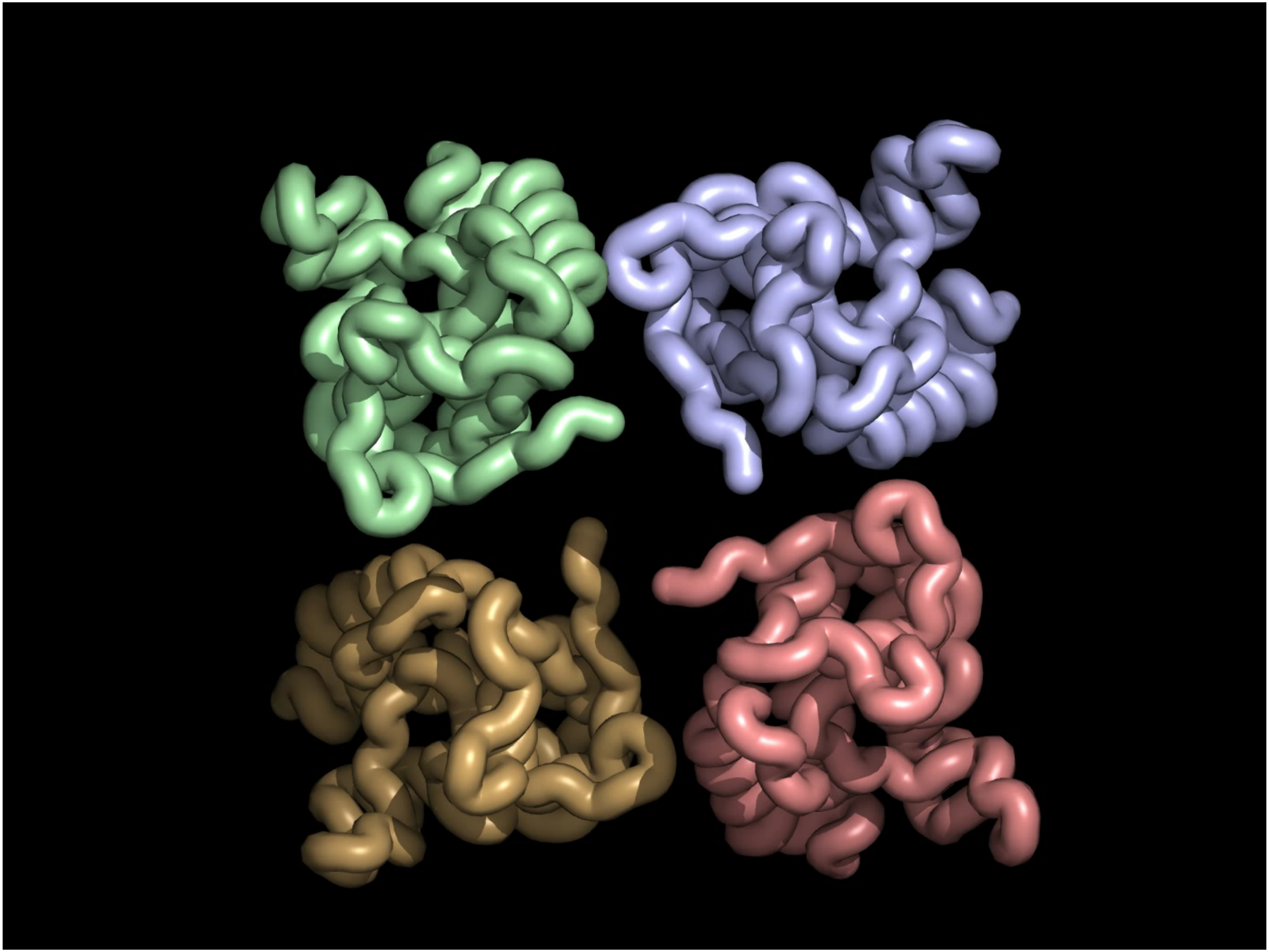




Staph Nuclease Simulation

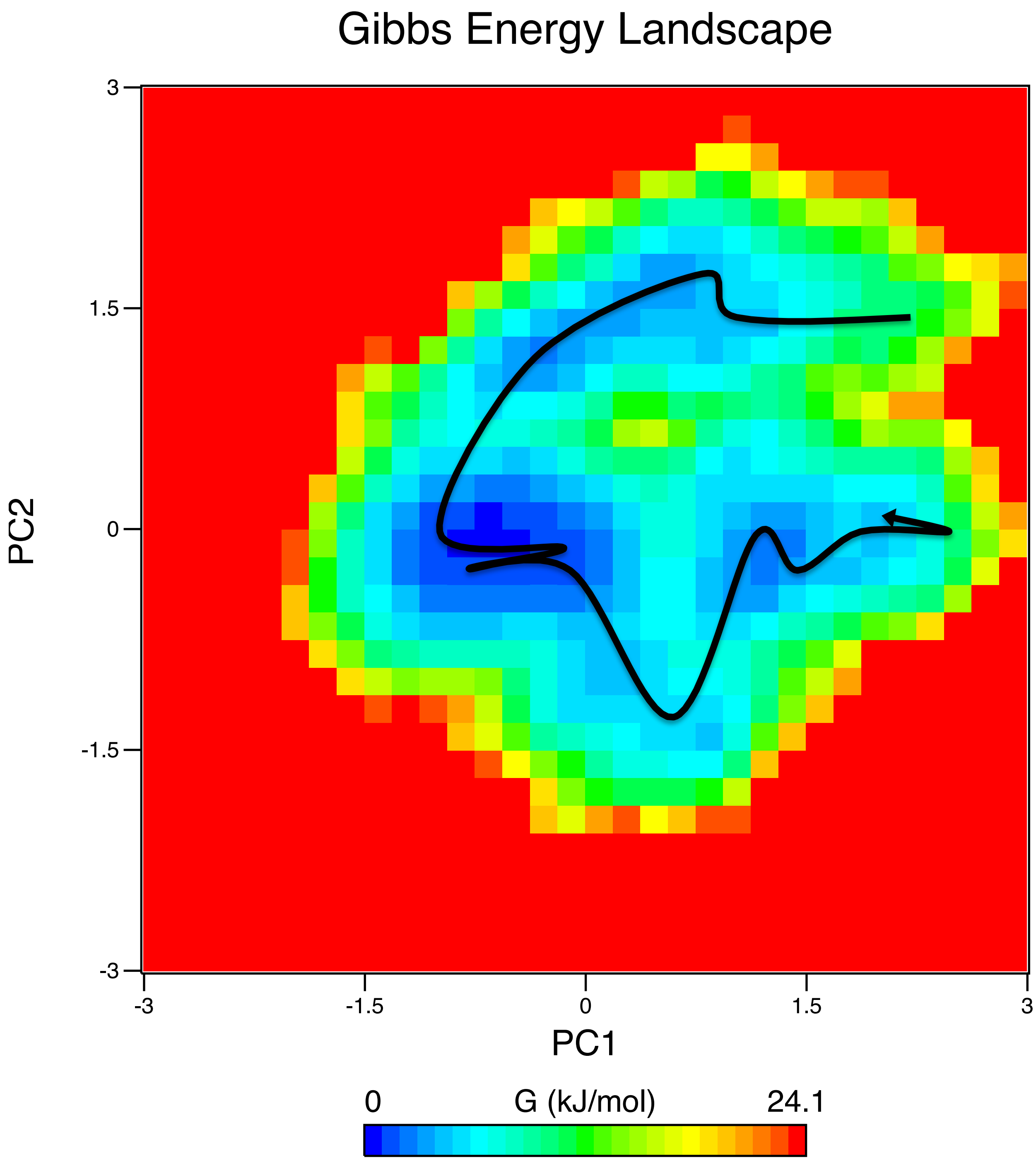
- PDB entry 1STN
- GROMACS, OPLSS-AA force field
- $P4_1$ unit cell
- 48.5x48.5x63.5
- 15,421 atoms
  - 8,904 protein
  - 6,477 TIP3P waters
  - 40 Cl^- ions
- T, P = const
- 1.1 microsecond



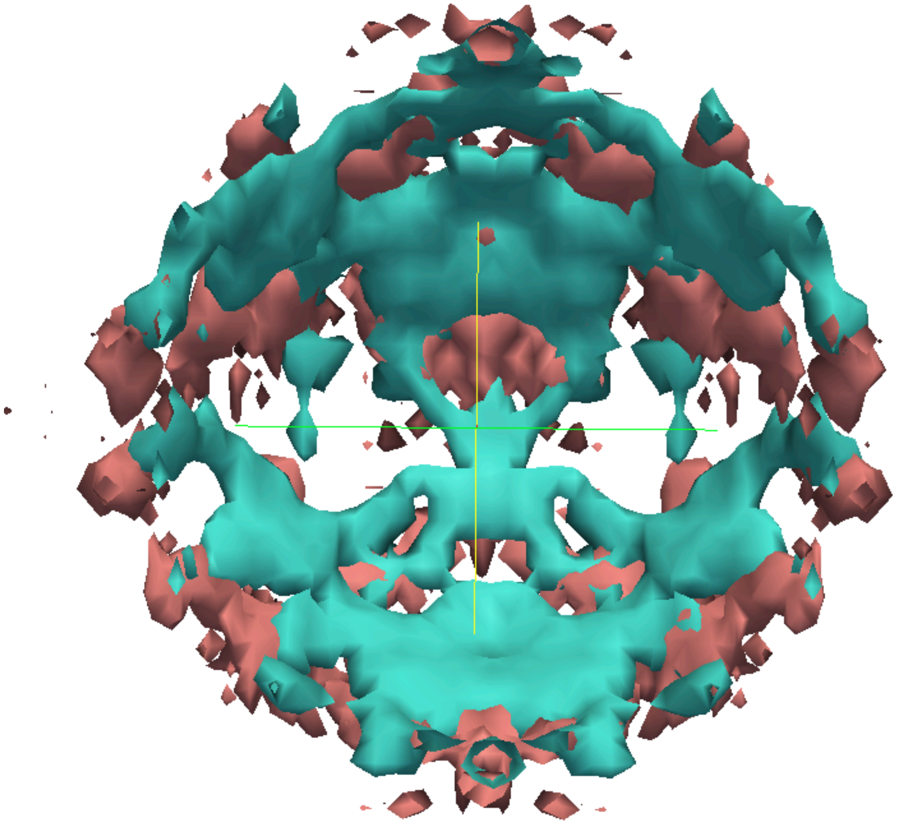


Combined Landscape

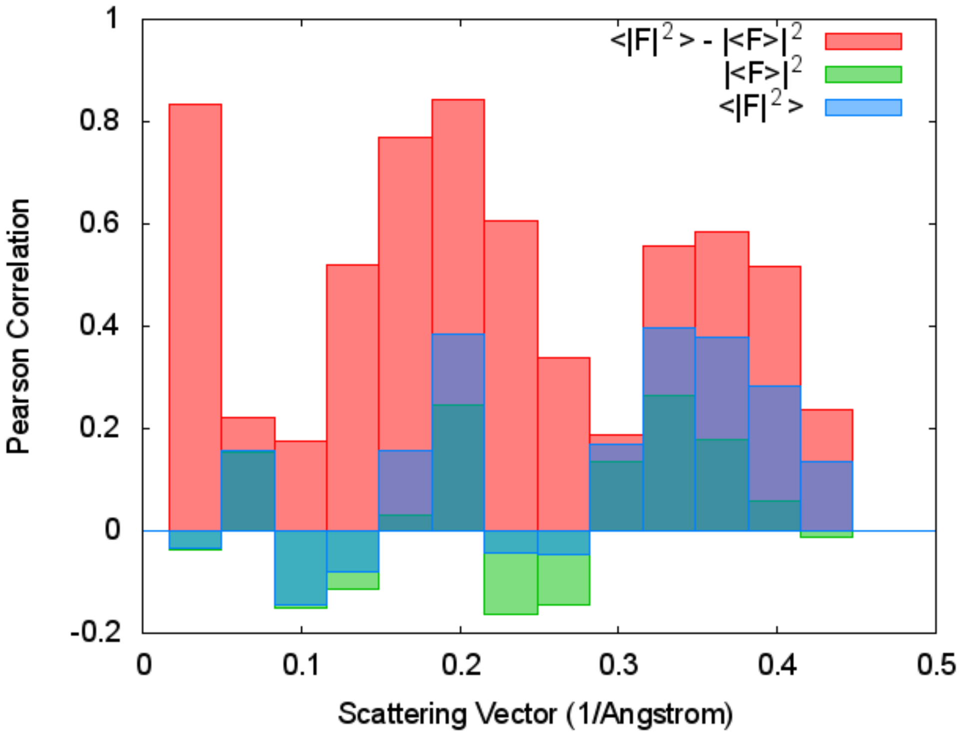
- 100-fold longer than previous simulation
- Trajectory explores several nearby basins



Comparison to Diffuse Data

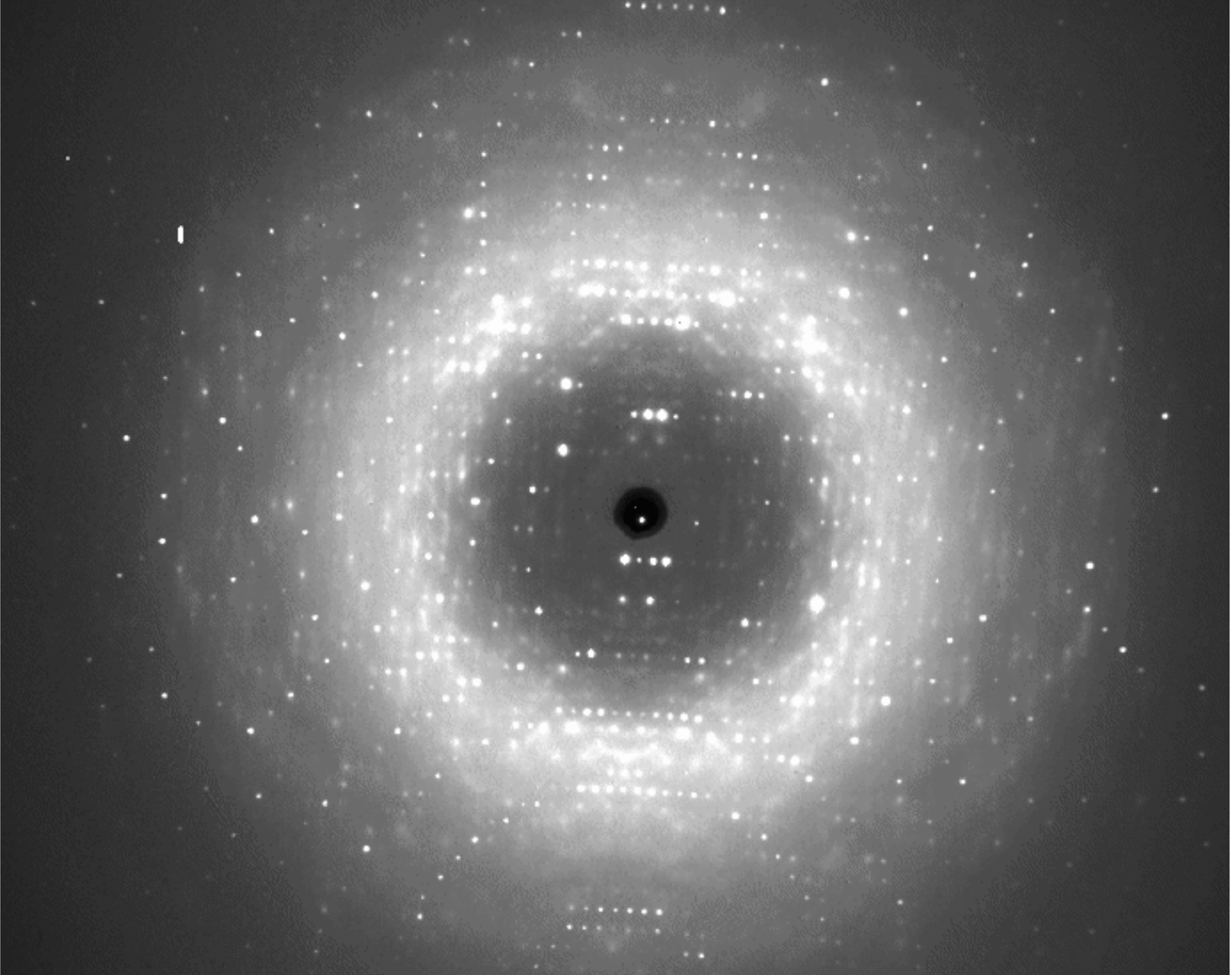


R = 0.066
C = 0.857



Anisotropic Correlations

Unpublished, with Tom Terwilliger, LANL



Diffuse Scattering ~~Can’t~~ Can Reveal Protein Dynamics

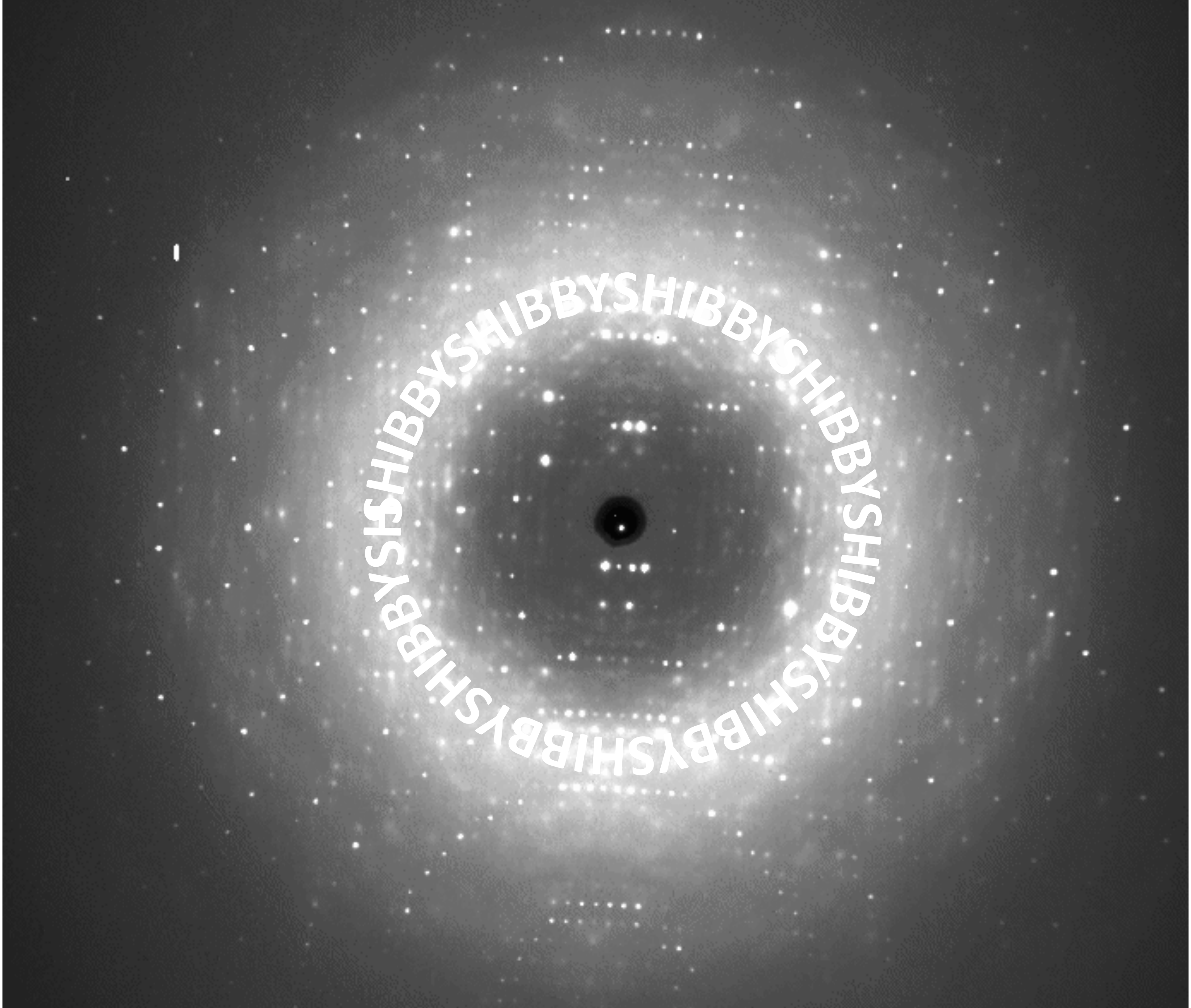
- Excellent agreement in isotropic component
- Convergence improved for 1 μs simulations
- Still challenges
  - Improve agreement in anisotropic component
    - Better Bragg and diffuse separation
  - Improve convergence
    - More sampling
    - Increased stability of simulation
  - Rigorously test assumptions

Real

Other Real Challenges

- The theory exists, but
  - It’s hard to understand
  - It’s even harder to communicate
  - It needs more work for macromolecular crystals
    - Multimodal atom position distributions
- Not part of the standard toolkit
- Few beamlines are set up for data collection
  - Minimize background sources
- Refinement algorithms are lacking
- Models don’t yet precisely match the information content of the signal
- Differences between the crystalline vs biological state

We can solve these problems



Acknowledgments

- Cornell
  - Sol Gruner
  - Steven Ealick
  - CHESS staff
- Rice
  - George Phillips, Jr.
  - James Clarage (now at St Thomas)
- FSU
  - Donald Caspar
- UCSF
  - James Fraser
  - Andrew Vanbenschoten
- LANL
  - Tom Terwilliger
- LBL
  - Nicholas Sauter
  - Paul Adams
- Funding
  - NSF, Welch Foundation, Keck Foundation, LANL LDRD