PROTEINS: Structure, Function, and Genetics 12:145–157 (1992)

# Correlations of Atomic Movements in Lysozyme Crystals

**James B. Clarage,[1] Michael S. Clarage,[1] Walter C. Phillips,[1] Robert M. Sweet,[2] and Donald L.D. Caspar[1]**
*[1]Rosenstiel Basic Medical Sciences Research Center and Department of Physics, Brandeis University, Waltham, Massachusetts 02254, and [2]Department of Biology, Brookhaven National Laboratory, Upton, New York 11973*

***ABSTRACT*** **Diffuse scattering data have been collected on two crystal forms of lysozyme, tetragonal and triclinic, using synchrotron radiation. The observed diffraction patterns were simulated using an exact theory for simple model crystals which relates the diffuse scattering intensity distribution to the amplitudes and correlations of atomic movements. Although the mean square displacements in the tetragonal form are twice that in the triclinic crystal, the predominent component of atomic movement in both crystals is accounted for by short-range coupled motions where displacement correlations decay exponentially as a function of atomic separation, with a relaxation distance of ≈ 6 Å. Lattice coupled movements with a correlation distance ≈ 50 Å account for only about 5–10% of the total atomic mean square displacements in the protein crystals. The results contradict various presumptions that the disorder in protein crystals can be modeled predominantly by elastic vibrations or rigid body movements.**

**Key words: thermal diffuse X-ray scattering, protein disorder, molecular dynamics**

## INTRODUCTION

Bragg reflections contain information about the average structure of a crystal, i.e., the average positions and mean square displacements of atoms, but no information about variations from the average. Variations in the structure of molecules in a crystal, most notably the correlations in atomic displacements, show up in the diffuse X-ray scattering. Analysis of the diffraction from tropomyosin crystals[1,32,33] showed that the predominant component of diffuse scattering can be accounted for by uncorrelated, large amplitude fluctuations (≈ 5 Å rms displacements) in the α-helical coiled coil strands in the crystal net, corresponding to gas-like coupling among the movements of the net segments, as in the Einstein model for a crystalline solid.[2] Streaks of diffuse scattering from orthorhombic lysozyme have been modeled in terms of rigid packing disorders in the lattice.[3] Our analysis of the diffuse X-ray scattering due to variations of the protein structure in insulin crystals demonstrated that the predominant component of the molecular displacements exhibits correlations over short distances roughly the size of amino acid residues.[4]

The intensities of Bragg reflections at the reciprocal lattice points of index (hkl) correspond to the squared Fourier transform of the average structure, $F^2(hkl) = |\mathrm{FT}\langle\rho(\mathbf{r})\rangle|^2$; whereas the total intensity scattered by the crystal, $F^2(\mathbf{R})$, as a function of the continuous reciprocal space coordinate $R = |\mathbf{R}| = 2 \sin\theta/\lambda$, is defined by the average of the squared Fourier transforms of the instantaneous structures, $\langle|\mathrm{FT}\rho(\mathbf{r})|^2\rangle$. Variational scattering corresponds to the difference of these two, that is the difference between the average of the square and the square of the average transform, or the variance in the transformed density. The variational scattering $I_V(\mathbf{R}) = \langle|\mathrm{FT}\rho(\mathbf{r})|^2\rangle - |\mathrm{FT}\langle\rho(\mathbf{r})\rangle|^2$ due to atomic movements with mean square displacement $\delta^2$ (in the direction of the vector $\mathbf{R}$) corresponds to the intensity lost from the Bragg reflections, which are damped as a function of reciprocal spacing by the Debye function $e^{-1/2BR^2} = e^{-(2\pi R\delta)^2}$. Therefore, the spherically averaged intensity of the variational scattering increases with reciprocal spacing relative to that of the Bragg reflections by the ratio $[1 - e^{-(2\pi R\delta)^2}]/e^{-(2\pi R\delta)^2} = [e^{(2\pi R\delta)^2} - 1]$. For example, if $\delta^2 = 0.25$ Å$^2$ (corresponding to a mean Debye temperature factor $B \approx 20$ Å$^2$, which is typical for many protein crystals) the integrated intensity of the diffuse scattering will exceed that of the Bragg diffraction beyond a resolution of $1/R = 3.8$ Å. Thus, at the resolution required to locate atoms in protein crystals, more X-rays are diffusely scattered than diffracted into Bragg reflections.

Modulations of $I_V(\mathbf{R})$ as a function of $\mathbf{R}$ in reciprocal space are inversely proportional to the distance in real space over which displacements are correlated. The variational scattering from rhombohedral insulin has two components[4]: a very diffuse part, which accounts for most of the variational scattering, corresponding to movements that are correlated over distances of only 6 Å; and sharper haloes

Received March 6, 1991; accepted May 28, 1991.
Address reprint requests to James B. Clarage at his present address: Department of Biochemistry and Cell Biology, P.O. Box 1892, Rice University, Houston, TX 77251.

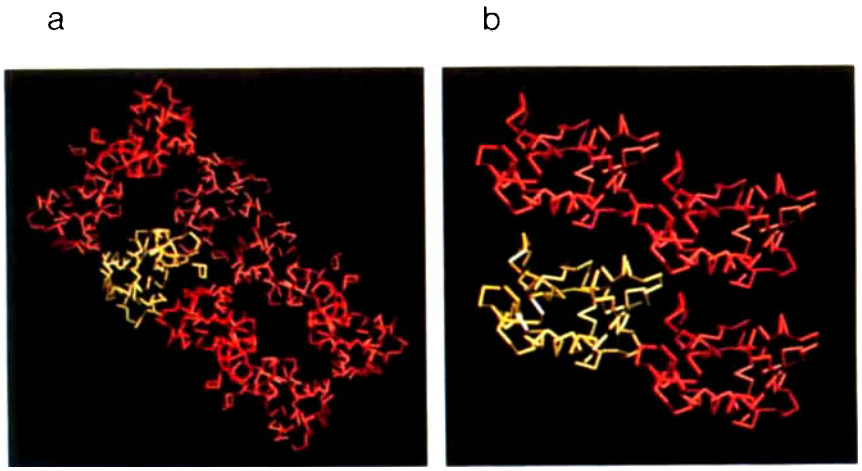


Fig. 1. Illustration of molecular packing in the two crystal forms studied. (**a**) View down the 4-fold axis of the tetragonal unit cell, which has eight protein molecules and 33.5% water by weight. (**b**) Four unit cells from the *a–b* plane of the more densely packed triclinic form containing 26% water.

around Bragg reflections that correspond to lattice coupled movements with correlations over distances the size of the protein molecules. Haloes from lattice coupled displacements are similar to those predicted by the Debye–Waller theory for crystals of elastically coupled rigid molecules.[5,6] The predominant very diffuse scattering corresponds to liquid-like coupling of internal movements.

Using Mössbauer crystallography, Nienhaus et al.[7] have established that the short-ranged coupled movements in myoglobin, giving rise to the very diffuse component of variational scattering, have relaxation times < 100 nsec. Thus these movements arise from thermal agitations—but these thermal fluctuations do not propagate as elastic lattice vibrations. The Mössbauer diffraction measurements also corroborate the inference from the diffuse scattering measurements on insulin[4] that the majority of the atomic disorder in crystalline proteins is due to intramolecular movements that are coupled only over short ranges. To determine how correlations in atomic movements depend upon the amplitudes of the displacements and the molecular packing, we have made synchrotron radiation measurements of the variational scattering from two different crystal forms of lysozyme: tetragonal and triclinic. The mean square atomic displacements in the tetragonal form are known from measurements of the temperature factors to be twice as large as in the more tightly packed triclinic lattice.[8,9] We have developed a simple analytical model based on the Patterson function for rigorous analysis of the relation between the variational intensity distribution and correlations in atomic displacement. This formalism enables exploration of different possible models for the disorder in the lysozyme crystals through computer simulation of the observed diffraction data.

## METHODS

### X-Ray Diffraction Data

#### *Data collection*

Hen egg white lysozyme was crystallized in two different space groups: triclinic ($P1$) and tetragonal ($P4_32_12$), according to standard procedures.[10] The tetragonal unit cell has $a = b = 79.2$ Å, $c = 38.0$ Å, and consists of 8 lysozyme molecules (Fig. 1a). The triclinic unit cell (Fig. 1b) contains one molecule, and has lattice parameters $a = 27.3$ Å, $b = 32.0$ Å, $c = 34.29$ Å; $\alpha = 88.5°$, $\beta = 108.6°$, $\gamma = 111.9°$.

Diffraction data for this study were collected on Kodak DEF X-ray film as still or small (0.5°) oscillation photographs at the Brookhaven National Synchrotron Light Source. The X-ray wavelength was 1.22 Å. With a flux of $10^9$ photons/sec through a 300 μm collimator, variational scattering features with optical densities of about 0.5 on the top film were recorded in 5–10 min. Since Bragg reflections could be detected within seconds, they are quite overexposed in these X-ray photographs.

#### *Data processing*

Optical densities from 0 to 2 on the films were digitized from 0 to 255 on a 100-μm raster. A pixel in the image corresponds to (1/700) $Å^{-1}$ in reciprocal space, and the diffracted beams cover 2–3 pixels. Thus, the order to order resolution of the recorded X-ray diffraction patterns is about 10 times the dimensions of a lysozyme molecule.

A global film fog was subtracted from the data, and the resulting intensities were corrected for the nonlinear response of the film and the capillary absorption. To remove the contributions from bulk water and Compton scattering, a circularly symmetric minimum background was subtracted from the data.

The background curve was defined by the average of the three or four lowest values at a given radius. This measurement of the spherically symmetric scattering may also contain contributions due to any completely random, gas-like component of the crystal disorder, which could be neglected since the focus was on analysis of the correlated movements.

The crystal orientation corresponding to each film was determined by finding the Ewald sphere orientation relative to the reciprocal lattice which gave the lowest mean square distance between the indexed spots and the sphere of reflection.

### *Data display*

Since the variational signal is relatively weak and diffuse, modulations in this signal are difficult to detect on a linear gray scale. Therefore, a color table was designed to highlight small differences in the distribution of the variational scattering, which visually breaks up into two categories: sharp haloes surrounding Bragg peaks, and more diffuse intensity spread throughout reciprocal space. On the byte scale of data, the diffuse component fills the range 0–40, which begins black and ends with cyan. Sitting atop this component are the sharper haloes extending from 40 to 125; this intensity range is coded from cyan to bright pink. Finally, the overexposed Bragg reflections, with intensity 255, are white in the color table (see Fig. 3).

## Theory of Variational Scattering

For clarity, the following discussion is restricted to one-dimensional systems, or equivalently, to higher dimensional systems in which the disorder is isotropic. The generalization to anisotropic disorder in higher dimensions is straightforward (see ref. 13 and Appendix). Formulas relating scattered intensity to atomic disorder will here be derived in terms of the average Patterson (autocorrelation) function $\langle P[\rho(r)]\rangle = \mathrm{FT}[I(R)] = \mathrm{FT}\,\langle|\mathrm{FT}\rho(r)|^2\rangle$, which is to be distinguished from the Patterson of the average structure $P[\langle\rho(r)\rangle] = \mathrm{FT}|\mathrm{FT}\langle\rho(r)\rangle|^2$. The widths of the peaks in $\langle P[\rho(r)]\rangle = \langle P(u)\rangle$ express the relative mean square displacements of atoms with a given pair separation, $u$. Thus, correlations in atomic displacements are manifest in how the widths of peaks in $\langle P(u)\rangle$ vary from those in the ideal (i.e., no motion) Patterson $P_0(u) = P[\rho_0(r)]$. Here, $\rho_0(r)$ is the density distribution with all atoms frozen at their most probable locations, and $\langle\rho(r)\rangle$ is the average density distribution. Presuming that all atoms have the same mean square fluctuation $\delta^2$ in their position, $\langle\rho(r)\rangle = \rho_0(r)*e^{-1/2(r/\delta)^2}$, where * denotes convolution.

In the Einstein model of a crystal, where units move independently, all peaks in $\langle P[\rho(r)]\rangle$ are identical to those in $P[\langle\rho(r)\rangle]$, except the origin peak which is always unbroadened (since in calculating the average Patterson at $u = 0$ the density has not been displaced from itself and is therefore in complete correlation). This abrupt change in the peak shapes from that of the ideal to that of the average reflects the fact that displacement correlations disappear as soon as one goes to a neighbor (Fig. 2a).

Two-dimensional (2-d) crystals of polystyrene latex spheres in water provide insight into another type of disorder.[11,12] These colloidal crystals originally caught out interest because of their model simplicity and direct visibility under a videomicroscope. Here, the average Patterson peaks *gradually,* monotonically broaden towards the limiting widths found in the Patterson of the average (Fig. 2b). This limiting width is approached exponentially,[12] the decay length being a measure of the range over which motions are correlated.

In these examples, the average Patterson $\langle P[\rho(r)]\rangle$ can be obtained by a suitable broadening of the ideal Patterson $P_0(u)$ toward the Patterson of the average $P[\langle\rho(r)\rangle]$, which has uniform peak widths of $2\delta^2$, the convolution of two density distributions with mean square displacements $\delta^2$. Explicitly, each peak in $P_0(u)$ is convoluted with a Gaussian whose squared halfwidth, or variance, $\sigma^2$ depends upon the Patterson coordinate $u$. Thus, correlations in the disorder for a crystal are governed by a variance function $\sigma^2(u)$, which we express as a fraction of the mean square displacement $\delta^2$ from the ideal lattice positions by writing $\sigma^2(u) = 2\delta^2[1 - \Gamma(u)]$. This is a reasonable form for $\sigma^2(u)$ because the width of the Patterson peak at $u$ gives the relative mean square displacement of atoms separated by $u$; that is, $\sigma^2(u) = \langle[\delta(r) - \delta(r + u)]^2\rangle_r$, where $\langle\ \rangle_r$ is an average over all lattice sites $r$. Multiplying out this expression for the relative mean square displacements puts it in the form $\sigma^2(u) = 2\delta^2[1 - \Gamma(u)]$, where $\Gamma(u) \equiv \langle\delta(r)\delta(r + u)\rangle_r/\delta^2$ is the displacement correlation function, which is unity at $u = 0$, where there is total correlation, and which vanishes for large separations where displacement correlations eventually disappear.

Fourier transforming a Patterson so broadened gives the average scattered intensity (see Appendix):

$$
\begin{aligned}
I(R) &= \mathrm{FT}[\langle P(u)\rangle] \\
&= e^{-(2\pi R\delta)^2} \sum_{n=0}^{\infty} \frac{(2\pi R\delta)^{2n}}{n!} \{F_0^2(R_{hkl}) * \mathrm{FT}[\Gamma^n(u)]\} \\
&= e^{-(2\pi R\delta)^2} F_0^2(R_{hkl}) \\
&\quad + e^{-(2\pi R\delta)^2} (2\pi R\delta)^2 \{F_0^2(R_{hkl}) * \tilde{\Gamma}(\Delta R)\} \\
&\quad + e^{-(2\pi R\delta)^2} \frac{(2\pi R\delta)^4}{2} \{F_0^2(R_{hkl}) * \tilde{\Gamma}(\Delta R) * \tilde{\Gamma}(\Delta R)\} \\
&\quad + \cdots \qquad (1)
\end{aligned}
$$

where $F_0^2(R_{hkl}) = \mathrm{FT}[P_0(u)]$ represents the set of ideal structure factors for the regularly ordered

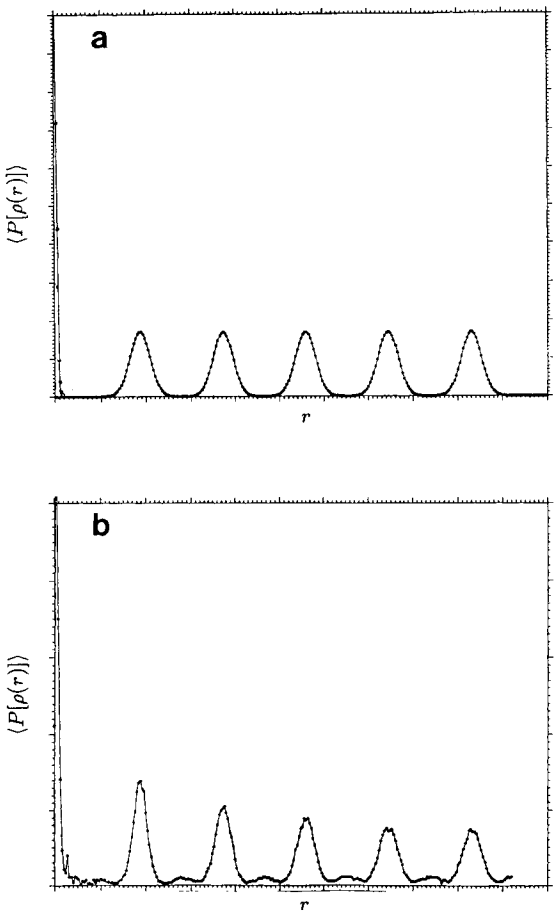


Fig. 2. The average Patterson function $\langle P[\rho(r)]\rangle$ (in arbitrary units) for (**a**) an Einstein crystal, and (**b**) the (1,1) lattice direction of a 2-d hexagonal crystal of latex spheres.[12]

structure; $\Gamma(\Delta R) \equiv \mathrm{FT}[\Gamma(u)]$ is defined as the halo function, that is, the Fourier transform of the displacement correlation function; and $\Delta R$ is a vector measured from each reciprocal lattice point $R_{hkl}$. The zeroth order term is the Bragg scattering, reflecting the average structure with uniform mean square displacements $\delta^2$ of the atoms from their ideal positions in the lattice, but nothing about correlations in the displacements since $\Gamma(u)$ is absent. Higher order terms correspond to variational scattering. For a given correlation function, $\Gamma(u)$, the equation represents the diffuse scattering as the sum of successive convolutions of the Bragg peaks with the halo function, $\bar{\Gamma}(\Delta R)$, with each term of order $n$ modulated by the Debye function $e^{-(2\pi R\delta)^2}$ and the factor $(2\pi R\delta)^{2n}/n!$.

In this representation of the variational scattering, the first-order term predominates if the displacements are small compared with the resolution. When the dimensionless expansion parameter $(2\pi R\delta) < 1$, higher order terms are relatively weak since they contain greater powers of $(2\pi R\delta)^2$, and since their signal is more spread out in reciprocal space due to multiple convolutions of the halo function. So for typical protein crystals with $\delta \approx 0.5$ Å, the first-order expression should be adequate out to resolutions of about 3 Å. This first-order term corresponds to the expression previously used for the variational scattering[4] that was based on an approximation of the average Patterson function.

At high resolution ($R > 1/2\pi\delta$), other terms in the power series contribute and the variational scattering becomes broader and broader due to the cumulative effects of successive convolutions of $\bar{\Gamma}(\Delta R)$. The extent of the diffuse scattering about the Bragg reflections is therefore a function of resolution. The limiting result of the summation is to give a featureless intensity distribution at high scattering angles, corresponding to the random, gas-like component of the motion; such a component necessarily exists, for without a random component to the correlation between nearest neighbors (i.e., a finite width in the first Patterson peak) all atoms would move in complete concert and the crystal would not be disordered.

Our simple theory assumes Gaussian probability

distributions for the atoms, and homogeneous disorder. The homogeneity assumption is inherent in a Patterson based formalism, for the Patterson function is sensitive only to relative atomic separations. The Gaussian assumption is key to analytically solving the Fourier integral of $\langle P(u)\rangle$, which gives the expression for the average scattered intensity (the Gaussian approximation is always justified to first order in the variational scattering[13]). The functional form chosen for $\sigma^2(u)$ implicitly assumes that our system is in fact a crystal, i.e., that the autocorrelation function does not diverge indefinitely; nonetheless, the formalism also applies to less ordered states of matter by choosing a suitably divergent variance function for the Patterson peaks.

If the total atomic mean square displacement $\delta^2$ can be factored into statistically independent components, then the halo function is the weighted sum of component halo functions, each describing correlations in one of the components of the motion. The relative weights of the component haloes are defined by the relative mean square displacements of the components. For instance, suppose the disorder is due to two distinct components (e.g., a long-range and a short-range correlated part) such that $\delta^2 = \delta_1^2 + \delta_2^2$. Then the first order variational scattering is (see appendix)

$$I_V(R) = e^{-[(2\pi R\delta_1)^2 + (2\pi R\delta_2)^2]} \times \{(2\pi R\delta_1)^2 F_0^2(R_{hkl}) * \mathrm{FT}[\Gamma_1(u)] + (2\pi R\delta_2)^2 F_0^2(R_{hkl}) * \mathrm{FT}[\Gamma_2(u)]\} \quad (2)$$

where $\Gamma_1(u)$ and $\Gamma_2(u)$ are the displacement correlation functions for the two components.

This treatment of the scattering, although motivated by the short-range exponential broadening of Patterson peaks [$\Gamma(u) = e^{-u/\gamma}$] seen in colloidal crystals, places no restrictions on the form of the correlation function $\Gamma(u)$. The equations accommodate any functional form for the correlation in atomic displacements, for example, the abrupt delta function broadening of the Einstein model[2] $\Gamma(u) = \Delta(u)$, or the long-range algebraic broadening of the Debye–Waller elastic model[5,6] $\Gamma(u) = 1/u$. [Since the Debye–Waller model predicts (see appendix) halos which fall off inversely with the square of the distance from the Bragg reflection, i.e., $\tilde{\Gamma}(\Delta R) \sim 1/\Delta R^2$, this corresponds to a correlation function of $\Gamma(u) = \mathrm{FT}[1/\Delta R^2] \sim 1/u$]. *All information about the nature of the correlations is contained in the halo function* $\tilde{\Gamma}(\Delta R)$, *which is in principle measurable.*

## Simulation

The Fourier transform of the halo function describes how displacements are coupled. Simple examination of an X-ray film can provide an estimate of the correlation scales by measuring the halo widths, which are inversely related to the distances over which movements are correlated. If the haloes are so broad that they overlap (implying that correlations extend only over distances smaller than the unit cell dimension), then the characteristic coupling distance is indicated by the width of modulations in the diffuse pattern, which can be larger than the spacing between neighboring Bragg reflections. To determine not only the length scale, but the functional form for the correlations, the shape of the halo must be measured. Unfortunately, this is not easy. The X-ray experiment only measures the intensity on the Ewald sphere; hence, only 2-d slices through the haloes are obtained. Even if the three-dimensional (3-d) intensity were mapped, the fact that haloes from neighboring peaks overlap makes deconvolution of the halo function an ill-posed problem, especially for the scattering due to intramolecular motion where the haloes overlap extensively to form a continuous diffuse intensity distribution.

The problems with extracting $\tilde{\Gamma}(\Delta\mathbf{R})$ from measurement of the diffuse intensity suggest the utility of applying convolution rather than deconvolution methods to analyze the correlation function. By assuming a form for the correlations, the X-ray data can be simulated. Unlike Bragg crystallography, where a subset of the X-ray film is integrated and reduced to a list of structure factors which are used to refine models for the average structure, this analysis proceeds by simulating all of the ($\approx 10^5$) pixels from a film as a means of refining the magnitude of the atomic displacements $\delta$, as well as the range and functional form of correlations in these displacements.

Previous simulations of insulin data[4] treated $I_V(\mathbf{R})$ as a 3-d array constructed by Fourier transformation of the ideal Patterson truncated with $\Gamma(\mathbf{u})$, rather than by the alternate procedure of convoluting the ideal Bragg peaks with the Fourier transform of $\Gamma(\mathbf{u})$. For the simulations of insulin data, the 3-d array was then sliced with an appropriately oriented Ewald sphere, and the intersected intensities stereographically projected onto a 2-d film array.

Instead of constructing the entire 3-d intensity, which involves calculating and storing many intensities not on the film, the present simulations began in film space. A point on the film array was mapped back to the corresponding coordinate $\mathbf{R}_E$ on the Ewald sphere, and the variational scattering $I_V(\mathbf{R}_E)$ was computed at this point to give the film intensity, after applying the inverse square law correction for the film geometry. Simulation of the first order variational scattering at $\mathbf{R}_E$ for a still photograph proceeds as follows: Given a trial correlation function $\Gamma(\mathbf{u})$ with characteristic decay length $\gamma$, we considered all reciprocal lattice nodes $\mathbf{R}_{hkl}$ within a distance of $\approx 1/\gamma$ of our sphere point $\mathbf{R}_E$, that is, all nodes close enough to affect the intensity at this point on the sphere of reflection. Then the contribution from each of these lattice nodes was summed at

$\mathbf{R}_E$, the weights being determined by the values of the halo functions $\bar{\Gamma}(\mathbf{R}_E - \mathbf{R}_{hkl})$, and the ideal structure factors $F_0^2(\mathbf{R}_{hkl})$. Structure factors for the ideally ordered crystal were computed from atomic coordinates for triclinic and tetragonal lysozyme[8,9] obtained from the Brookhaven protein data bank.[14] Lastly, the convolution products $F_0^2(\mathbf{R}_{hkl})*\bar{\Gamma}(\mathbf{R}_E - \mathbf{R}_{hkl})$ were multiplied by $e^{-(2\pi R_E\delta)^2}(2\pi R_E\delta)^2$, where $\delta$ is the estimate for the rms displacement. Higher order terms in the variational scattering use a similar algorithm. Oscillation photographs were simulated by superposing several still simulations contained within the oscillation range.

Since this simulation method does not rely upon actually constructing the 3-d distribution of $I_V(\mathbf{R})$ in the computer, but rather performing operations on precise points in reciprocal space, one can explore arbitrarily fine reciprocal spacings on the film, i.e., arbitrarily long length scales in the crystal lattice. To simulate the 20 pixels between two Bragg peaks in the triclinic data, for instance, it is not necessary to construct, store, and perform calculations on an actual 3-d model consisting of 20 unit cells on a side, which would be computationally prohibitive. The only arrays involved are the 2-d film array and the list of ideal structure factors. This procedure also makes it possible to simulate the Bragg scattering, $e^{-(2\pi R_{hkl}\delta)^2}F_0^2(\mathbf{R}_{hkl})$, for those reciprocal lattice points that lie within the beam divergence of the Ewald sphere, by treating the X-ray beam profile, estimated from the low angle Bragg reflections, as a sharp Gaussian halo function centered at each reciprocal lattice point.

## RESULTS

### X-Ray Diffraction Data

Several still and oscillation diffraction patterns, spanning a 90° orientation range, were collected from both crystal forms of lysozyme. The center column of Figure 3 shows three processed photographs chosen for quantitative analysis by the simulation procedure. The first (Fig. 3a) is a still photograph from triclinic lysozyme, oriented 17.9° from (010). The other two are 0.5° oscillation photographs from tetragonal lysozyme, separated by 90° from each other; one (Fig. 3b) is oriented 12.7° from the 4-fold (001) axis, and the other (Fig. 3c) is 13.4° from (110). These films are representative of the data set. Since the triclinic unit cell is nearly isodimensional, the overall intensity distribution from these crystals looks very similar in different orientations. The lattice sampling in the tetragonal form, however, is quite different parallel and perpendicular to the 4-fold axis since the $c^*$-axis is twice as large as $a^*$ and $b^*$, a fact which accounts for the different appearance of the two views shown in Figure 3b and c. Data taken from different crystals in the same orientations demonstrated the reproducibility of the diffuse scattering features in the diffraction patterns.

In the three photographs (Fig. 3a–c), displayed to a resolution of 3 Å, the Bragg peaks are immediately evident as overexposed (intensity 255 on this 0–255 scale) regularly spaced spots; their inherent size can be judged from the lowest order reflections. With increasing resolution, variational scattering between the Bragg peaks becomes visible, more so in the tetragonal data, indicating greater atomic motion in these crystals. The non-Bragg intensity visually separates into two components: one concentrated around the Bragg reflections and another more spread out through reciprocal space. The sharper haloes (intensity 40–200) indicate a component of the motion coupled through the lattice, while the broader scattering (intensity 0–40) is due to movements whose correlations decay in a distance much less than the unit cell dimension, so that the haloes of neighboring Bragg reflections overlap to produce a quite diffuse intensity distribution.

Although the typical signal in the very diffuse scattering is approximately three orders of magnitude weaker than the strong Bragg reflections, as judged from short exposures, the diffuse signal is spread throughout reciprocal space whereas the Bragg signal is not. For instance, in the triclinic crystal, the order to order resolution of $\approx$ 10 times the lattice constant means that the Bragg scattering is confined to but a thousandth of the reciprocal lattice volume. Hence at the resolution at which the spherically averaged diffuse scattering and Bragg intensities are equal ($1/R \approx 3$ Å), the average intensity of the diffuse scattering is $\approx 10^{-3}$ that of the mean Bragg peak. In addition, even though pixel for pixel the intensity of the broad component of the variational scattering is much less than the sharp component, integrations of the data from insulin[4] and lysozyme show that 80–95% of the total variational intensity resides in the very diffuse component, indicating that most of the disorder is due to movements correlated over short range.

### Simulation of Diffraction Patterns

At first glance, it is not obvious that all the variational scattering can be simulated by a convolution procedure, let alone using the same halo function for all reciprocal lattice points. For example, in the triclinic photograph there are examples at the same resolution of a small intense spot next to a large weak spot, seemingly implying two different halo functions. However, this effect is generally due to a Bragg spot with a structure factor too small to produce a significant halo lying directly on the Ewald sphere, and another Bragg spot far from the sphere, yet with such an intense structure factor that the associated halo extends far enough into reciprocal space to intersect the Ewald sphere. Indeed, in the still diffraction pattern of the triclinic crystal, there are only a handful of reciprocal lattice points actually in reflecting position and therefore most of the

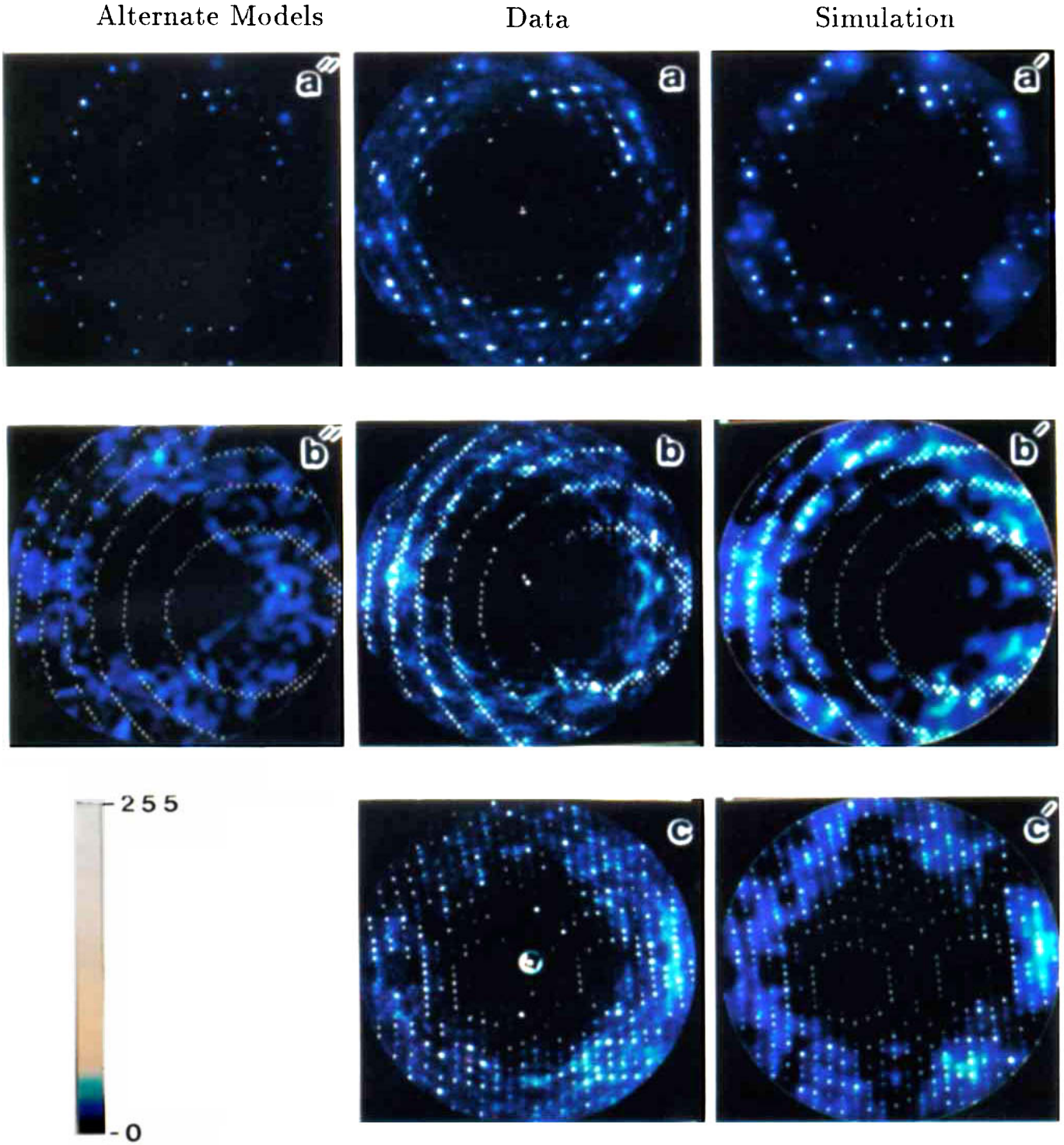


Fig. 3. Results of the analysis of diffuse synchrotron X-ray scattering from tetragonal and triclinic lysozyme crystals. The center column shows the processed film data, displayed out to 3 Å resolution. The Bragg reflections are overexposed. (**a**) Still photo of triclinic lysozyme, with X-ray beam directed 17.9° from (010). (**b**) Small oscillation (0.5°) photo of tetragonal lysozyme, oriented 12.7° from the 4-fold (001) axis. The spindle axis is vertical in this figure. (**c**) A 0.5° oscillation photo of the same crystal as in (**b**), but rotated 90° about the spindle, such that the view is 13.4° to (110). In the right column are simulations of the Bragg and diffuse scattering data shown in the center column, assuming a liquid-like model for the movements in which displacement correlations decay exponentially with atomic separation. (**a'**) Simulation of the triclinic data, using two components to account for the atomic displacement. The long-range component, which gives rise to sharp haloes surrounding Bragg reflections, has correlation length $\gamma_L$ = 50 Å and mean square displacement $\delta_L^2$ = (0.11 Å)$^2$ = 0.012 Å$^2$. The short-range component, responsible for the broader intensity distributed between peaks, has correlation length $\gamma_S$ = 6 Å and mean square displacement $\delta_S^2$ = (0.33 Å)$^2$ = 0.11 Å$^2$. Short-range correlated movements account for 90% of the total mean square displacement $\delta^2 = \delta_S^2 + \delta_L^2$ = (0.35 Å)$^2$ = 0.12 Å$^2$. (**b'**) and (**c'**) show simulations of the tetragonal data using two components. The long-range component has correlation length $\gamma_L$ = 50 Å and mean square displacement $\delta_L^2$ = (0.11 Å)$^2$ = 0.012 Å$^2$, while the short-range component has $\gamma_S$ = 6 Å and mean square displacement $\delta_S^2$ = (0.49 Å)$^2$ = 0.24 Å$^2$. Short-range movements account for 95% of the total mean square displacement $\delta^2$ = (0.50 Å)$^2$ = 0.25 Å$^2$. The left column shows simulations of the diffraction data in the center column, assuming alternate models for the coupling of atomic displacements. (**a''**) A simulation treating the triclinic protein crystal as an elastic solid, with thermal waves giving rise to haloes around the Bragg reflections whose intensity falls off with the square of the distance from reciprocal lattice points. (**b''**) Simulation of the view near the 4-fold of tetragonal lysozyme, in which the protein molecule is assumed to undergo a hinge-bending mode.

sharp spots are just the central parts of haloes around lattice points close to the Ewald sphere.

Other features in the data appear to bear no relation to the reciprocal lattice, and therefore seem incapable of arising from haloes centered on Bragg reflections. For example, in the view near the 4-fold axis of tetragonal lysozyme (Fig. 3b), the circle of peaks due to the intersection of the basal plane ($hk0$) with the Ewald sphere has a horseshoe shaped intensity sitting inside, with no apparent connection to the reciprocal lattice. However, this feature is actually due to the superposition of haloes around Bragg reflections in the ($hk\bar{1}$) plane, reflections which are themselves invisible in this orientation. The tetragonal data display more of these sorts of features than the triclinic patterns since the smaller tetragonal reciprocal lattice dimensions make it more likely for haloes to overlap.

The features in the lysozyme variational scattering data were reasonably well simulated as a sum of contributions from two component halo functions $\tilde{\Gamma}_S$ and $\tilde{\Gamma}_L$, each corresponding to spherically symmetric fluctuations whose displacement correlations decay exponentially over short- and long-range coupling distances $\gamma_S$ and $\gamma_L$, respectively. The mathematical form of the halo functions $\tilde{\Gamma}(\Delta R)$ is $\mathrm{FT}[e^{-u/\gamma}] = 8\pi\gamma^3/[1+(2\pi\Delta R\gamma)^2]^2$, which is a three-dimensional analog of a Lorentzian.[13] The variational scattering from the lysozyme crystals was approximated as

$$I_V(\mathbf{R}) = e^{-[(2\pi R\delta_S)^2+(2\pi R\delta_L)^2]} \times \{(2\pi R\delta_S)^2 F_0^2(\mathbf{R}_{hkl}) * \mathrm{FT}[e^{-u/\gamma_S}] + (2\pi R\delta_L)^2 F_0^2(\mathbf{R}_{hkl}) * \mathrm{FT}[e^{-u/\gamma_L}]\} \quad (3)$$

where displacements with rms amplitude $\delta_S$ have a characteristic short-range decay length of $\gamma_S$, and longer-range correlated movements with amplitude $\delta_L$ decay over a distance $\gamma_L$. The results of the simulations are shown in Figure 3a′–c′. The halo function $\mathrm{FT}[e^{-u/\gamma_S}]$ spreads intensity much farther from the Bragg reflections than in the Debye–Waller model,[6] yet not to the point of producing a spherically symmetric diffuse scattering pattern as in the Einstein model.[2,15] The short-range decay of exponential displacement correlations lies between the idealized extremes of long-range algebraic decay for elastic coupling and abrupt delta function decay for uncorrelated fluctuations.

### Measurement of $\delta_S$, $\gamma_S$ and $\delta_L$, $\gamma_L$

A rough estimate for the rms displacement of the short-range correlated component of motion gives $\delta_S \approx 0.5$ Å for tetragonal and $\delta_S \approx 0.3$ Å for triclinic. Thus, the first-order expression for $I_V(\mathbf{R})$ was deemed sufficient to simulate the most interesting features in our diffuse scattering data, which occur at a resolutions of around $1/R = 4$ Å, for in this case $(2\pi R\delta) < 1$. The estimate for $\delta_S$ comes from measuring the radial distribution of the circularly averaged variational intensity,[4] which is governed by $f(R)[1-e^{-(2\pi R\delta_S)^2}]$, where $f(R)$ is the mean atomic form factor for carbon, nitrogen, and oxygen. To estimate the second order correction, note that the halo will be due to two powers of $\Gamma(u)$ and will therefore have an effective decay length of $\gamma/2$. Since the halo function scales with the cube of the decay length, the second-order halo will have peak values 1/8 those of the first-order halo. Combining this factor with the extra factor of $(2\pi R\delta)^2/2$ in the second-order expression makes this term $\approx 4\%$ as intense as the first order term at $1/R \approx 4$ Å.

With the initial estimates for $\delta_S$, trial and error simulations of the short-range component of the variational scattering were performed to fit for $\gamma_S$. For all three films, $\gamma_S = 6$ Å gave the best fit (Fig. 3a′–c′). The fringes in the simulated intensity distributions are sensitive to changes in $\gamma_S$ of about an Angstrom. Smaller values produce intensity which is too circularly symmetric, and values much larger produce too much modulation. Once the correlation length was set, the original estimate for the rms displacement was refined. The resulting rms displacements for the two crystal forms are $\delta_S = 0.49$ Å for tetragonal and $\delta_S = 0.33$ Å for triclinic, corresponding to mean square displacements of 0.24 and 0.11 Å$^2$, respectively. These mean squared values are estimated to be reliable to $\approx 10\%$.

After obtaining $\delta_S$ and $\gamma_S$, the parameters for the long-range component were dealt with. Again, an initial estimate was used for the rms displacement. As previously mentioned, $\delta_L < \delta_S$ because there is less integrated intensity in the sharp haloes. Different values of $\gamma_L$ were then explored until the simulated halo widths matched those in the data. All three films were best fit with $\gamma_L = 50$ Å. This value is reliable to about 5 Å. A subsequent refinement of $\delta_L$ to scale the ratio of the long-range component of the variational scattering relative to the short-range component, in accord with the experimental data, leads to rms displacement amplitudes of $\delta_L = 0.11$ Å for both crystal forms.

The total mean square atomic displacements $\delta^2 \equiv \delta_S^2 + \delta_L^2$ for tetragonal and triclinic lysozyme are then $\delta^2 = 0.25$ Å$^2$ and $\delta^2 = 0.12$ Å$^2$, respectively, a result in excellent agreement with measurements of the mean temperature factors of these crystals from the fall-off in the intensity of the Bragg scattering[8,9] (the mean square displacements computed from the temperature factors in the protein data bank files[14] are $\delta^2 = 0.25$ Å$^2$ for tetragonal and $\delta^2 = 0.13$ Å$^2$ for triclinic). From the diffuse scattering measurements the relative contributions to the total mean square displacements from the short range coupled motions (a quantity which cannot be factored from temperature factors) is $\delta_S^2/\delta^2 = 0.95$ for tetragonal and $\delta_S^2/\delta^2 = 0.90$ for triclinic. Furthermore, the contribution of any completely uncorrelated atomic movements to the crystallographic temperature fac-

tors must be negligible over the resolution range of our measurements because our exclusion of any spherically symmetric diffuse scattering did not diminish the agreement with the total mean square displacements measured from the Bragg intensities.

### Alternate Models for Displacement Correlations

In order to check whether the movements in the protein crystal could be approximated by an elastic solid model, an attempt was made to simulate the data using the $1/\Delta R^2$ halo function of the Debye–Waller theory. [In doing so, a subtlety arises as to how to scale such a simulation to other models, for the halo function $1/\Delta R^2$ cannot be normalized. To normalize, we impose a highest spatial frequency $1/d$ corresponding to the shortest wavelength disturbance in the solid. In this case, the normalized halo function is[13] $\tilde{\Gamma}(\Delta R) = (d/4\pi)(1/\Delta R^2)$; $\Delta R \leq 1/d$. The smallest measurable value of $\Delta R$ is set by the divergence of the X-ray beam.] Figure 3a″ shows the results from simulating the variational scattering from triclinic lysozyme using the $1/\Delta R^2$ halo function with a Debye wavelength cutoff of $d = 3$ Å, corresponding to the mean noncovalent distance of closest interatomic approach in the crystal. This model gives a poor fit to the data, concentrating too much intensity close to the reciprocal lattice points, thereby failing to account for the gradual fall-off in halo intensity around Bragg reflections and the very diffuse distribution between peaks.

Simulations were also performed with a simple hinge model for the protein movements by directly computing $I_V(\mathbf{R}) = \langle|\mathrm{FT}\rho(\mathbf{r})|^2\rangle - |\mathrm{FT}\langle\rho(\mathbf{r})\rangle|^2$ over a series of conformations in which two lobes of the molecule oscillate about a hinge axis.[16] The results are shown in Figure 3b″. Such models produce relatively high spatial frequency modulations in the calculated diffuse intensity, characteristic of the dimensions of the rigid domains which are significantly larger than the reciprocal of the experimentally measured diffuse fringe widths. Thus, any model with rigid body movements of large domains (corresponding to low frequency normal modes) would predict more modulation than experimentally observed in the diffuse scattering patterns. In contrast, models with short-range homogeneous correlations in the atomic displacements provide a reasonable simulation of the diffuse scattering observed from tetragonal and triclinic lysozyme crystals.

## DISCUSSION

The atomic disorder in tetragonal and triclinic lysozyme crystals can be accounted for in terms of two components. Simulations of variational scattering data indicate that the principal component of atomic movement, responsible for 90–95% of the mean square displacement, has short-range correlations that decay over a relaxation distance of $\approx 6$ Å, which is the size of an amino acid or a few atomic coordination shells. The minor component of the mean square displacement is due to movements of protein molecules that are correlated over distances larger than the molecular dimensions. The halo function which best fits the data for the predominant diffuse component corresponds to an exponential displacement correlation function, i.e., $\Gamma(u) = e^{-u/\gamma}$. Thus, the observed displacement correlations are liquid-like in the sense that they decay exponentially with the distance between atoms, just as *positional* correlations in a liquid.

Crystal contacts are known to influence atomic mobility.[17] The differences in disorder between the tetragonal and triclinic forms of lysozyme reflect the effect of crystal interactions on the atomic displacements. The more tightly packed triclinic molecule has half the mean square displacement of the tetragonal form. Our factorization of the total mean square displacement shows that both forms, however, have the same magnitude for the long-range lattice coupled contribution; therefore, the extra mobility in tetragonal lysozyme is due to increased intramolecular movements compared to triclinic lysozyme. The looser tetragonal environment does not significantly affect the lattice dynamics but does liberate internal motions not accessible to the triclinic form.

Our results show that the major contribution to temperature factors in protein crystals is due to the short-range correlated, liquid-like displacements that give rise to the predominant component of the diffuse X-ray scattering. This conclusion is at odds with surmises based on normal mode calculations[19] and models with rigid body movements,[20] as well as inferences from measurements of Young's modulus[21] and the speed of sound[22] in protein crystals, all of which implicate long-range coupled displacements as the major contribution to temperature factors. These models with long wavelength vibrations or rigid body movements predict amplitudes of displacement that can be fit to the amplitudes measured from crystallographic temperature factors but no attempts have been made to compare the correlations in the model movements with experimental data.

Given the order to order resolution of our recorded diffraction patterns, the movements included in this study are correlated over distances less than $\approx 10$ molecular lengths. X-Ray crystallographic measurements of Bragg reflections integrate out the contribution of disorder due to lattice vibrations with wavelengths larger than the reciprocal of the spot size.[18] Thus, the integrated peak intensities provide information only about fluctuations that are correlated over distances less than the reciprocal of the Bragg spot dimensions. Scattering due to any fluctuations correlated over distances larger than the reciprocal of the beam divergence will appear as

part of the Bragg intensity. Because we can only measure the diffuse scattering beyond the limit of the Bragg spots, our analysis applies to the same component of motion contained in crystallographic B-values. Indeed, the total rms atomic displacements determined from measurements of the variational scattering are in excellent agreement with measurements of the temperature factors from these crystals.[8,9]

Unlike ordinary X-ray diffraction, Mössbauer diffraction[23] is able to separate the inelastic from the elastic scattering at the reciprocal lattice nodes and therefore provides a direct measure of the magnitude of motion due to the longer wavelength lattice fluctuations. Using this technique, Nienhaus et al.[7] have concluded that the total contribution from lattice coupled movements with relaxation times $< 100$ nanoseconds in myoglobin crystals is $\delta^2 = 0.02$ Å$^2$, which is only a small fraction of the intramolecular mean square displacements in these protein crystals. Most of the inelastic scattering measured by Mössbauer diffraction is very diffusely scattered[7] corresponding to the diffuse intensity distribution we have mapped between the Bragg reflections in other protein crystals.

Our X-ray measurements have excellent momentum resolution and can therefore give accurate estimates of the distance scales involved in atomic movements. Mössbauer crystallography, due to its energy resolution, can resolve times scales up to 100 nsec,[7] and has been used to establish the dynamic behavior of the short-range coupled motions in protein molecules. These Mössbauer studies also shed light on why lattice correlated motions appear to contribute such a small fraction of the total mean square displacement $\delta^2$ in protein crystals. The Debye elastic theory predicts a linear relation between $\delta^2$ and absolute temperature. Above 200 K, however, this scaling has been observed to break down for the iron atom[24] in myoglobin, and displacements grow much faster as a function of temperature. Inelastic neutron scattering results[25] have confirmed this phenomenon for the entire protein molecule. Since linearly extrapolating the low temperature scaling of $\delta^2$ to room temperature gives a value in agreement with the lattice contribution to motion, it seems that the non-Debye disorder is due to the excitation of the liquid-like intramolecular modes.

The protein crystal appears to behave as an elastic solid below some critical temperature, with the mean square displacements growing linearly with temperature and displacement correlations presumably decaying algebraically with distance. However, above $\approx 200$ K there is a transition to the liquid-like intramolecular motions we observe at room temperature. Such a transition from long-range algebraic to short-range exponential correlation in the *displacements* is analogous, at least mathematically, to the phase transition in certain spin systems[26] where the *spin* correlations change from algebraic to exponential decay above some critical temperature, or in the phase transition from solid to hexatic in 2-d crystals of hard disks, where the *positional* correlations undergo this change.[27] Exponential displacement correlations are observed in colloidal crystals[12] where the ordering is due to repulsive interactions among loosely packed particles; thus these dynamic systems may provide a better analogy for movements in crystals of protein molecules than the lattice vibrations of elastic crystals.

Our conclusions are based upon simulations of diffraction data using theoretical methods developed for relating correlations in atomic displacement to variational scattering. The advantage of this simple theory—which points out that the Fourier transform of the variational halo function is the displacement correlation function—is that it presumes no a priori form for atomic correlations (such as Debye normal modes). A limitation is the implicit assumption of homogeneous disorder. The disorder in protein crystals does appear fairly homogeneous, however, as evidenced by how well the X-ray data can be simulated by our simple Patterson space model. There is little contribution to the mean square displacements from large-scale movements of specific rigid domains. Rather, the atomic displacements are dominated by local movements similar to the conformational adaptability proteins exhibit in different crystal environments.[28] Such homogeneity is probably due to the fact that globular proteins have a fairly uniform arrangement of atoms and the noncovalent, intramolecular interactions are of the same type as those between neighboring proteins within the crystal. Nonetheless, there are undoubtedly inhomogeneities in protein disorder, and the next step, as previously suggested,[4] is to test more detailed models for protein motion (e.g., from molecular dynamics calculations[29]) againt variational scattering data by directly computing $I_V(\mathbf{R}) = \langle|\mathrm{FT}\rho(\mathbf{r})|^2\rangle - |\mathrm{FT}\langle\rho(\mathbf{r})\rangle|^2$ from a set of model conformations.

## APPENDIX: THEORY OF VARIATIONAL SCATTERING IN TERMS OF THE PATTERSON FUNCTION*

For simplicity, the reasoning begins with point atoms in 1-d; later the generalization to higher dimensions is discussed. As outlined in the text, the average Patterson can be derived from the ideal one by smearing the peaks in the latter, each by a particular (normalized) Gaussian. This variable width Gaussian has variance of the form $\sigma^2(u) = 2\delta^2[1 - \Gamma(u)]$, where $\delta^2$ is the mean square fluctuation in the position of an atom from its ideal lattice point, and $\Gamma(u)$ is the function which governs how the displace-

*By James B. Clarage.

ments of neighboring atoms are correlated. The variable smearing of the ideal Patterson translates into placing a Gaussian, with halfwidth $\sigma^2(u_j)$, at each Patterson peak $P_o(u_j)$, where $j$ labels the Patterson peaks. Thus, the average Patterson is

$$\langle P(u)\rangle = \sum_j P_o(u_j)[2\pi\sigma^2(u_j)]^{-1/2} e^{-1/2(u-u_j)^2/\sigma^2(u_j)}.$$

The corresponding intensity as a function of the scattering vector $k = 2\pi R$ is given by Fourier transformation.

$$I(k) = \mathrm{FT}[\langle P(u)\rangle]$$

$$= \int du e^{iku} \sum_j P_o(u_j)[2\pi\sigma^2(u_j)]^{-1/2} e^{-1/2(u-u_j)^2/\sigma^2(u_j)}.$$

Pulling a change of variable, $v = u - u_j$, will allow partial evaluation of the transform,

$$I(k) = \int dv \sum_j e^{iku_j} P_o(u_j) e^{ikv} [2\pi\sigma^2(u_j)]^{-1/2} e^{-1/2v^2/\sigma^2(u_j)}.$$

Performing the $v$ integration,

$$I(k) = \sum_j e^{iku_j} P_o(u_j) e^{-1/2k^2\sigma^2(u_j)}$$

$$= \mathrm{FT}[P_o(u) e^{-1/2k^2\sigma^2(u)}].$$

Writing this in terms of $\Gamma(u)$ gives

$$I(k) = e^{-(k\delta)^2} \mathrm{FT}[P_o(u)\, e^{(k\delta)^2\Gamma(u)}]\,. \qquad (4)$$

The expression for the scattered intensity in formula (4) is most readily interpreted by expanding the exponential inside the transform in power series,

$$I(k) = e^{-(k\delta)^2} \sum_{n=0}^{\infty} \frac{(k\delta)^{2n}}{n!} \mathrm{FT}[P_o(u)\Gamma^n(u)] \qquad (5)$$

or,

$$I(k) = e^{-(k\delta)^2} \sum_{n=0}^{\infty} \frac{(k\delta)^{2n}}{n!} \{F_o^2(k_{hkl}) * \mathrm{FT}[\Gamma^n(u)]\}$$

$$= e^{-(k\delta)^2} F_o^2(k_{hkl})$$

$$+ e^{-(k\delta)^2} (k\delta)^2 \{F_o^2(k_{hkl}) * \tilde{\Gamma}(\Delta k)\}$$

$$+ e^{-(k\delta)^2} \frac{(k\delta)^4}{2} \{F_o^2(k_{hkl}) * \tilde{\Gamma}(\Delta k) * \tilde{\Gamma}(\Delta k)\}$$

$$+ \cdots$$

where $F_0^2(k_{hkl}) = \mathrm{FT}[P_o(u)]$ represents the set of ideal structure factors for the regularly ordered structure, and $\tilde{\Gamma}(\Delta k) = \mathrm{FT}[\Gamma(u)]$.

The Fourier transform of the displacement correlation function is the halo function $\tilde{\Gamma}(\Delta k)$, which determines the diffuse intensity distribution at a distance $\Delta k$ from each Bragg peak at $k_{hkl}$. The higher order terms representing more diffuse components of the variational scattering correspond to successive convolutions of the halo function with itself. Represention of the diffuse scattering in terms of the halo function can be thought of as an application of the Wiener–Kintchine theorem,[30] which states that for a random, homogeneous function the power spectrum is the Fourier transform of the autocorrelation function. Interpreting the diffuse scattering surrounding a Bragg peak as a collection of Rowland ghost spectra due to periodic displacements in the crystal,[31] the halo is the power spectrum of spatial frequencies comprising the displacements, and is therefore the Fourier transform of the displacement spatial-autocorrelation function. That is, the halo function equals $\mathrm{FT}[\Gamma(u)]$.

### Components of Disorder

Often, it is useful to characterize the disorder in a system in terms of separate components. For example, if the motion were due to two distinct components (e.g., a long-range and a short-range component as in the case of a protein crystal), then the squared-halfwidths of the Patterson peaks would be given by the statistical sum of two terms

$$\sigma^2(u) = \sigma_1^2(u) + \sigma_2^2(u)$$
$$= 2\delta_1^2[1 - \Gamma_1(u)] + 2\delta_2^2[1 - \Gamma_2(u)]\,.$$

The corresponding intensity is then a generalization of Eq. (4)

$$I(k) = e^{-[(k\delta_1)^2 + (k\delta_2)^2]} \mathrm{FT}[P_o(u)\, e^{(k\delta_1)^2\Gamma_1(u) + (k\delta_2)^2\Gamma_2(u)}]$$

and the first order variational scattering is

$$I_V(k) = e^{-[(k\delta_1)^2 + (k\delta_2)^2]} \times \{(k\delta_1)^2 F_0^2(k_{hkl}) * \mathrm{FT}[\Gamma_1(u)] + (k\delta_2)^2 F_0^2(k_{hkl}) * \mathrm{FT}[\Gamma_2(u)]\}\,. \qquad (6)$$

If there are many independent components, then the total $\Gamma(u)$ is simply the sum over all the component correlation functions. For example, consider the Debye–Waller model of an elastic solid.[5,6] Here the motion is due to an ensemble of normal modes, each with a particular wavelength $\Lambda$, which are assumed to be statistically independent of each other (random phase relations). Each mode, with wavenumber $q = 2\pi/\Lambda$ contributes to the Patterson by periodically broadening the peaks, $\sigma_q^2(u) = 2\sigma_q^2[1 - \cos(qu)]$. The total correlation function is then the weighted sum of the individual ones $\Gamma_q(u) = \cos(qu)$. Since the energy in a mode is proportional to the square of the frequency $\omega(q)$, equiparition implies that the relative amplitudes of the modes scale as $1/\omega^2(q)$ which for acoustic modes in an elastic solid [$\omega(q) \propto q$] is the same as $1/q^2$. So,

$$\Gamma(u) = \sum_q \frac{1}{q^2}\cos(qu)$$

$$= \mathrm{FT}\left[\frac{1}{q^2}\right].$$

Thus, in three dimensions $\Gamma(u) \sim 1/u$, so correlations in displacement decay algebraically. The halo function is $\tilde{\Gamma}(\Delta R) = 1/\Delta R^2$, and thus the first-order diffuse scattering intensity will fall off inversely with the square of the distance from the Bragg reflection.

### Anisotropy in Higher Dimensional Systems

The arguements above were restricted to 1-d homogeneous systems, or equivalently to higher dimensional systems which are homogeneous and isotropic. The assumption of homogeneity is inherent in a Patterson based formalism; however, it is possible to treat the anisotropic coupling of atomic movements which can arise in higher dimensional systems.[13] In more than one dimension replace $(k\delta)^2$ by $(\mathbf{k}\cdot\boldsymbol{\delta}_j)^2$ since the scattering vector $\mathbf{k}$ can resolve motion parallel only to its own direction. $\boldsymbol{\delta}_j$ is a vector with the magnitude and direction of the $j$th component of motion in the system. The expression for the scattering is then

$$I(\mathbf{k}) = e^{-\sum_j(\mathbf{k}\cdot\boldsymbol{\delta}_j)^2}\,\mathrm{FT}[P_o(\mathbf{u})\exp\{\sum_j(\mathbf{k}\cdot\boldsymbol{\delta}_j)^2\Gamma_j(\mathbf{u})\}]$$

where the summations run over the independent components of motion. $\Gamma_j(\mathbf{u})$ is the function which describes how atomic displacements along $\boldsymbol{\delta}_j$ are correlated throughout the crystal, and if it is an anisotropic function (e.g., a higher dimensional analog to an exponential, with different correlation lengths in different directions of space) it will give rise to an anisotropic halo function $\tilde{\Gamma}_j(\Delta\mathbf{k}) \equiv \mathrm{FT}[\Gamma_j(\mathbf{u})]$ surrounding the Bragg peaks. The term out in front, $e^{-\sum_j(\mathbf{k}\cdot\boldsymbol{\delta}_j)^2}$, is the generalized Debye function. Expanding the intensity expression in power series gives

$$\begin{aligned} I(\mathbf{k}) = {} & e^{-\sum_j(\mathbf{k}\cdot\boldsymbol{\delta}_j)^2}F_o^2(\mathbf{k}_{hkl}) \\ & + e^{-\sum_j(\mathbf{k}\cdot\boldsymbol{\delta}_j)^2}\sum_j(\mathbf{k}\cdot\boldsymbol{\delta}_j)^2\{F_o^2(\mathbf{k}_{hkl}) * \tilde{\Gamma}_j(\Delta\mathbf{k})\} \\ & + e^{-\sum_j(\mathbf{k}\cdot\boldsymbol{\delta}_{j'})^2}\sum_j\sum_{j'}(\mathbf{k}\cdot\boldsymbol{\delta}_j)^2(\mathbf{k}\cdot\boldsymbol{\delta}_{j'})^2 \\ & \{F_o^2(\mathbf{k}_{hkl}) * \tilde{\Gamma}_j(\Delta\mathbf{k}) * \tilde{\Gamma}_{j'}(\Delta\mathbf{k})\} \\ & + \cdots \end{aligned} \tag{7}$$

Again, the zero order term is the Bragg scattering, the rest variational scattering. To lowest order we see that the variational scattering from statistically independent components is simply given by the sum of the scattering from the individual components. However, higher order terms include crossing between the components.


## ACKNOWLEDGMENTS

Diffraction data were collected at Brookhaven National Laboratory in the Biology Department single-crystal diffraction facility of the National Synchrotron Light Source, which is supported by the United States Department of Energy. This work was supported by U.S. Public Health Service Grant CA47439 from the National Cancer Institute to DLDC.


## REFERENCES


1. Boylan D., Phillips, G.N. Motions of tropomyosin. Biophys. J. 49:76–78, 1986.
2. Einstein, A. Die Planck'sche Theorie der Strahlung und die Theorie der Spezifischen Wärmen. Ann. Phys. 22:180–190, 1907.
3. Doucet, J., Benoit, J.P. Molecular dynamics studied by analysis of the x-ray diffuse scattering from lysozyme crystals. Nature (London) 325:643–646, 1987.
4. Caspar, D.L.D., Clarage, J., Salunke, D.M., Clarage, M. Liquid-like movements in crystalline insulin. Nature (London) 332:659–662, 1988.
5. Debye, P. Zur Theorie der Spezifischen Wärmen. Ann. Physik 39:789–839, 1912.
6. Waller, I. Zur Frage der Einwirkung der Wärmebewegung auf die Interferenz von Röntgenstrahlen. Z. Phys. 17:398–408, 1923.
7. Nienhaus, G.H., Heinzl, J., Huenges, E., Parak, F. Protein Crystal Dynamics studied by time resolved analysis of X-ray diffuse scattering. Nature (London) 338:665–666, 1989.
8. Blake, C.C.F., Koenig, D.F., Mair, G.A., North, A.C.T., Phillips, D.C., Sarma, V.R. Structure of hen egg-white lysozyme. Nature (London) 206:757–761, 1965.
9. Moult, J., Yonath, A., Traub, W., Smilansky, A., Podjarny, A., Rabinovich, D., Saya, A. The structure of triclinic lysozyme at 2.5 Å resolution. J. Mol. Biol. 100:179–195, 1976.
10. Steinrauf, L.K. Preliminary x-ray data for some new crystalline forms of β-lactoglobulin and hen egg-white lysozyme., Acta Crystallogr. 12:77–79, 1959.
11. Murray, C.A., Van Winkle, D.H. Experimental Observation of two-stage melting in a classical two-dimensional screened Coulomb system. Phys. Rev. Lett. 58:1200–1203, 1987.
12. Clarage, M., Clarage, J., Caspar, D.L.D. Correlated movements in two dimensional crystals. In preparation.
13. Clarage, J. Disorder in protein crystals. Doctoral Thesis, Brandeis Univ., 1989.
14. Bernstein, T.F., et al. The Protein Data Bank: A computer-based archival file for macromolecular structures. J. Mol. Biol. 112:535–42, 1977.
15. Debye, P. Interferenz von Röntgenstrahlen und Wärmebewegung. Ann. Phy. 43:49–95, 1914.
16. McCammon, J.A., Gelin, B.R., Karplus, M., Wolynes, P.G. The hinge-bending mode in lysozyme. Nature (London) 262:325–326, 1975.
17. Phillips, G.N. Comparison of the dynamics of myoglobin in different crystal forms. Biophys. J. 57:381–383, 1990.
18. Cochran, W. The correction of measured structure factors for thermal diffuse scattering. Acta Crystallogr. A25:95–101, 1969.
19. Diamond, R. On the use of normal modes in thermal parameter refinement: Theory and application to the bovine pancreatic trypsin inhibitor. Acta Crystallogr. A46:425–435, 1990.
20. Kuriyan, J., Weis, W. I. Rigid protein motion as a model for crystallographic temperature factors. Proc. Natl. Acad. Sci. U.S.A. 88:2773–2777, 1991.
21. Morozov, V.N., Morozova, T.Y.A. Thermal motion of whole protein molecules in protein solids. J. Theor. Biol. 121: 73–88, 1986.
22. Edwards, C., Palmer, S.B., Emsley, P., Helliwell, J.R., Glover, I.D., Harris, G.W., Moss, D.S. Thermal motion in protein crystals estimated using laser-generated ultra-

sound and Young's modulus measurements. Acta Crystallogr. A46:315–320, 1990.
23. O'Connor, D.A., Butt, N.M. The detection of the inelastic scattering of gamma rays at crystal diffraction maxima using the Mössbauer effect. Phys. Lett. 7:233–235, 1963.
24. Parak, F., Knapp, E.W., Kucheida, D. Protein dynamics: Mössbauer spectroscopy on deoxymyoglobin crystals. J. Mol. Biol. 161:177–194, 1982.
25. Doster, W., Cusack, S., Petry, W. Dynamical transition of myoglobin revealed by inelastic neutron scattering. Nature (London) 337:754–756, 1989.
26. Kogut, J.B. An introduction to lattice gauge theory and spin systems. Rev. Mod. Phys. 51:659–713, 1979.
27. Nelson, D.R., Halperin, B.I. Dislocation-mediated melting in two dimensions. Phys. Rev. B19:2457–2484, 1979.
28. Badger, J. Altered conformations of porcine insulin caused by crystal packing interactions. Submitted.
29. Karplus, M., McCammon, J.A. Dynamics of proteins: Elements and function. Annu. Rev. Biochem. 53:263–300, 1983.
30. Wiener, N. Generalized harmonic analysis. Acta Math. 55: 117–258, 1930.
31. James, R.W. "The Optical Principles of the Diffraction of X-rays." London: G. Bell, 1948.
32. Chacko, S., Phillips, G.N. Diffuse X-ray scattering from tropomyosin Crystals. Biophys. J. (in press).
33. Chacko, S. Doctoral Thesis, Univ. of Illinois, 1991.