Diffuse x-ray scattering from tropomyosin crystals

Susan Chacko*‡ and George N. Phillips, Jr.‡
*Department of Physiology and Biophysics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801; and ‡Department of Biochemistry and Cell Biology and the W.M. Keck Center for Computational Biology, Rice University, Houston, Texas 77005

ABSTRACT Diffuse scattering analyses are emerging as a technique to extract additional dynamic information from x-ray diffraction data. In fact, when examined carefully, most protein crystals show significant diffuse scattering in addition to the usual Bragg diffraction. This diffuse scattering contains information about the disorder in the crystal that cannot be obtained from the Bragg diffraction data. Diffraction from tropomyosin crystals shows characteristic diffuse scattering streaks that are directly related to motion of the molecules. The structure of tropomyosin to 15 Å resolution shows that the limited molecular contacts between molecules allow large conformational fluctuations of up to 8 Å amplitude. Models for the three-dimensional motion of tropomyosin have been tested by comparing their predicted diffuse scattering patterns with the experimental data. From the parameters of the successful simulations, we were able to determine the amplitudes, directions, and distances over which the atomic displacements are correlated.

INTRODUCTION

Classical x-ray crystallographic techniques offer a wealth of detailed spatial information about the static average structures of proteins. Even in crystalline form, however, proteins have a wide range of motions. These motions have been observed by isotope-exchange experiments, flash photolysis, nuclear magnetic resonance (NMR), neutron scattering, and catalytic activity measurements as well as x-ray diffraction. It is difficult, however, to separate static from dynamic displacements using x-ray diffraction, because the information about the displacements in the x-ray data is averaged over time and over all the unit cells in the crystal. Any kind of atomic or molecular displacement within the crystal causes a general decrease in the intensities of the Bragg diffraction spots, which can be analyzed using the Debye-Waller B-factors to approximate simple isotropic or anisotropic amplitudes of atomic displacements. In cases where the B-factors are sensitive to temperature in a reversible way, it can be inferred that the displacements are, indeed, dynamic in nature (1, 2). The Bragg data analysis of B-factors, however, does not provide any information about the correlation of the motions within the crystal.

With the decrease in intensity of the Bragg spots that accompanies motions, there is a concomitant increase in the intensity of the regions between Bragg spots. This non-Bragg scattering is commonly referred to as diffuse scattering, because it usually takes the form of diffuse haloes around the Bragg spots or a very diffuse cloud spread over a large area of reciprocal space. These two types of diffuse scattering have distinct causes; motions correlated over distances that are short compared to the unit cell dimensions give rise to the very diffuse cloud, while motions correlated over distances of the order of the unit cell or greater give rise to haloes around the Bragg spots. Caspar (3) has recently recommended using the term "variational" scattering for the diffraction effects of short and long-range correlated motions, since the long-range part is not particularly diffuse.

The effects of static lattice distortions and imperfections on the x-ray data must also be considered when analyzing crystallographic data. Classical x-ray diffraction does not distinguish between molecules in motion and those which are displaced from their ideal lattice position due to static lattice effects. These lattice effects, of course, have no biological significance in most cases, and are therefore not as interesting as the dynamic aspects of the displacements. Because the energy for intramolecular motion is thermal energy, decreasing the temperature should cause a decrease in the amplitudes of motion of the atoms, while the effects of lattice distortion remain unchanged. Thus, a study of the temperature dependence of the calculated B-factors provides an estimate of the amount of lattice distortion. Phillips et al. (4) reported a pronounced increase in resolution and intensity with decreasing temperature for tropomyosin, implying that most of the disorder is thermal in nature in these crystals. Frauenfelder et al. (1) used Mössbauer data, which is insensitive to static lattice disorder, coupled with x-ray data to estimate the

Address correspondence to Dr. Phillips.

0006-3495/92/05/1256/11 $2.00

lattice contribution to total disorder in myoglobin. At room temperature, they estimated the static lattice disorder to be less than 30% of the total disorder for myoglobin. Petsko and Ringe (5) modified this number in light of recent theories about the Mössbauer effect to be less than 15% for most crystals that diffract to high resolution.

The effects of motion within the crystal were originally examined by Debye (6) and Einstein (7) who showed that for simple lattices and independently moving atoms the Bragg data was superimposed on a monotonically increasing diffuse scattering function. With coupled motions between atoms, each crystal reflection is superimposed on a broad diffuse scattering maximum, and the width and height of these maxima increase with scattering angles. Because most crystals have a combination of coupled and uncoupled motions, both the monotonic and the Bragg-coupled features are seen in the diffuse scattering. B-factors can be theoretically calculated from the maximum frequency of elastic vibration of the crystal (8) but a similar calculation of the diffuse scattering, even for simple cubic crystals, requires a detailed knowledge of all the frequencies of vibration. For most molecular crystals, this calculation is immensely complex and not feasible in practice. The qualitative picture of the diffuse scattering, however, remains much the same for all crystals. The diffuse scattering from crystals of small molecules as well as macromolecules has both the Bragg-related and very diffuse scattering features described above.

Diffuse scattering has been observed in many different protein and nucleic acid crystals, (e.g. lysozyme (9, 10), insulin (3), myoglobin and myohemerythrin (G. N. Phillips, unpublished observations), 6-phosphogluconate dehydrogenase (11), tRNA (12), DNA (13) and tropomyosin (4)) but for the crystallographer who wants to determine a high-resolution structure it is an undesirable effect, complicating calculation of the background intensity. Only recently has some work focused on obtaining useful information from the diffuse scattering streaks and clouds. One can obtain information about the direction and amplitude of the atomic displacements, the distances over which they are coupled, and lattice substitutions. Another reason to study diffuse scattering is that the long-range correlated displacements give rise to haloes around the Bragg spots: these haloes blend with the main spot and affect its shape, which may affect the accuracy of measurement of the spot intensity. Hence, some understanding of the effects of molecular disorder is important even for accurate structure determination using classical x-ray crystallography.

Phillips et al. (4) previously reported diffuse scattering streaks of two types in the scattering from tropomyosin crystals. They observed large diffuse clouds along the main scattering directions, and smaller streaks associated with the Bragg spots perpendicular to the clouds. These features were interpreted as being evidence of considerable disorder in the crystal. Indeed, this crystal is 95% solvent, and the molecules are long filaments with little intermolecular contact. Hence, it is quite reasonable that there should be large fluctuations of the filament perpendicular to its axis, leading to diffuse scattering patterns of both the short-range as well as the long-range coupled types.

Further work on tropomyosin was carried out by Boylan and Phillips (14). They recorded the diffuse scattering for the 100 view of the crystal. They then successfully simulated the diffuse scattering in projection by modeling the motions of this filamentous molecule as being similar to a vibrating string. They obtained values for the amplitudes of motion that were consistent with those obtained from the B-factor analysis of tropomyosin.

Tropomyosin is especially suitable for analysis of dynamics, as the crystals produce strong, distinctive diffuse scattering patterns. The structure of tropomyosin in the Bailey crystal form (15) showed that the two polypeptide chains of the molecule are in exact register with their side chains interlocking, and they wind along the body diagonal of the unit cell in an elliptical path (see Fig. 1 *c*, *inset*). There is a nine-residue overlap at the head-to-tail region, where molecules join in forming filaments. Each molecule interacts with the others at two regions and is relatively free from interaction between those regions. The two regions between cross-over points are usually referred to as the long arm (COOH-terminal half of the molecule) and the short arm (NH_2-terminal half of the molecule). The head-to-tail overlap region is in the middle of the short arm. In the crystal, the filament makes half a turn along its length while in the muscle it makes a full turn, but the structures in crystal and muscle are otherwise quite similar.

There is much evidence for filament motion in the crystal. Previous studies have shown that some regions of the molecule are fluctuating perpendicular to the axis of the filament with large amplitudes of up to 8 Å (14). The regions of high interaction (cross-over points) have very little flexibility, the short arm region moves with smaller amplitudes than the long arm, and both regions move anisotropically with respect to the filament axis. Diffraction data at different temperatures (4) show that the intensity scattered by the long arm region is highly temperature-sensitive, with the Bragg intensities decreasing markedly at higher temperatures. This is an indication that the disorder is dynamic rather than static in nature. Hence there is considerable evidence to show

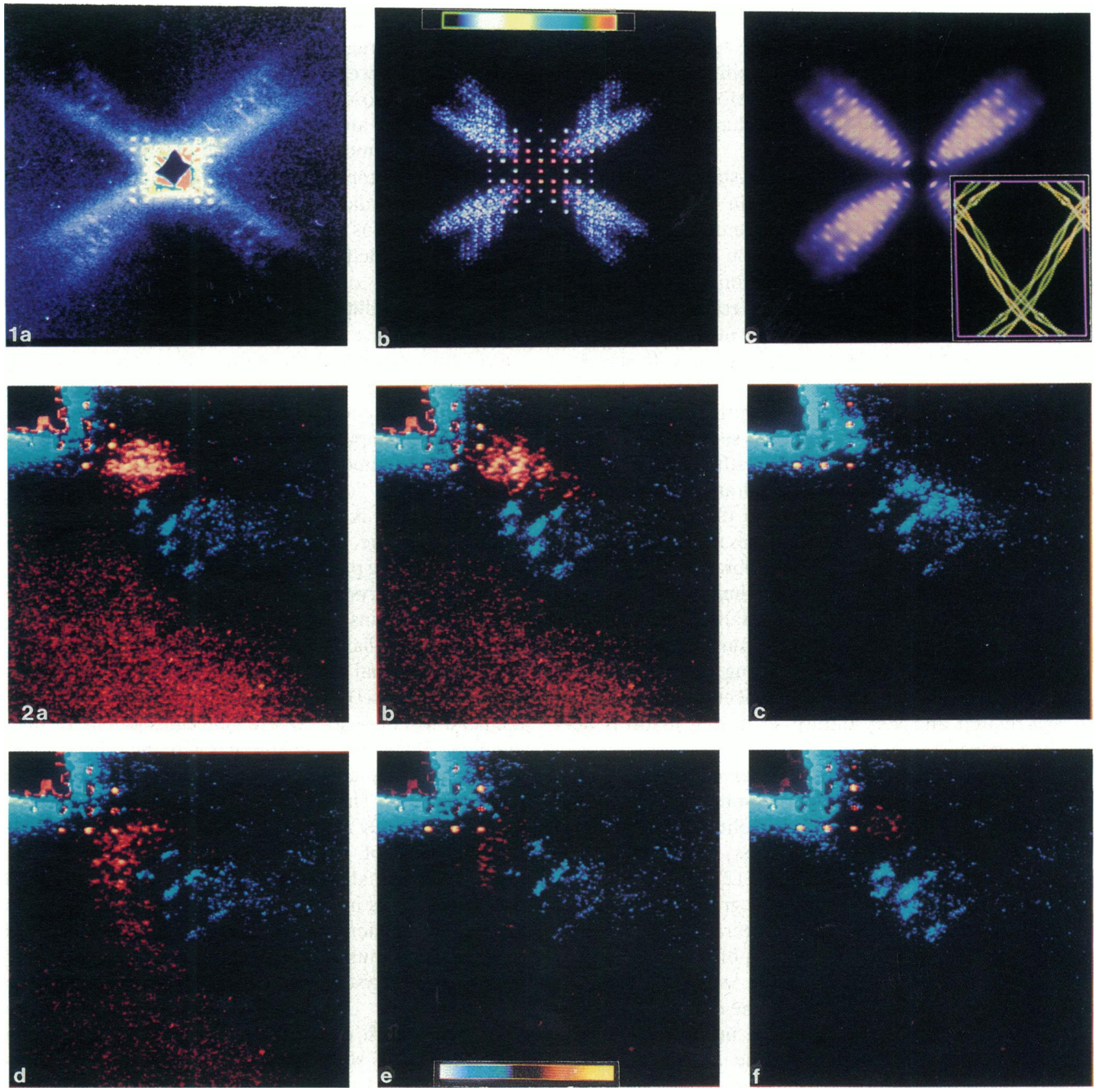


FIGURE 1 Diffuse scattering photographs and simulations of the 100 view. (*a*) Experimental data. (*b*) Simulation by superlattice method (*c*) Simulation by analytical method. (*Inset*) Corresponding projection of the molecular structure showing the arrangement of filaments.

FIGURE 2 "Difference" maps created by pixel-to-pixel subtraction of the experimental from the simulated data. The positive points (simulation greater than film intensity) are colored red-to-yellow, while negative points (film intensity greater than simulation) are colored blue-to-white. The better simulations have blacker difference maps. These simulations were performed using the superlattice method, and only the short-range correlated motions, giving rise to the large diffuse clouds, were simulated. Root mean square amplitudes of motion for the simulations were as follows: (*a*) Long arm 9.6 Å, short arm 9.6 Å. (*b*) Long arm 9.6 Å, short arm 4.4 Å. (*c*) Long arm 4.4 Å, short arm 4.4 Å. (*d*) Long arm 6.5 Å, short arm 6.5 Å. (*e*) Long arm 6.5 Å, short arm 4.4 Å. (*f*) Long arm 6.5 Å, short arm 2.4 Å. The best simulation was determined to be (*e*).

that the fluctuations are quite significant at physiological temperatures.

In this study, the diffuse scattering from the Bailey crystals was recorded on film and digitized. It was then simulated according to two different approaches; the methods of Phillips et al. (14) extended to a full three-dimensional simulation, as well as the more theoretical approach of Caspar and co-workers (3, 10, 16), also in three dimensions. Although the experimental data does not include a full three-dimensional diffuse scattering data set, a number of sections of the simulated data were compared to the available experimental data.

METHODS

Tropomyosin was prepared as described by White (17) and crystallized in capillaries in the Bailey crystal form (18) as described by Phillips (19). The experimental data were collected on a rotating anode generator with a double-mirror focusing system. A helium tunnel was used between the crystal and film to minimize air scatter. The data were collected at about 4°C by using a jet of cold air to cool the crystal. The data consisted of still photographs which were exposed for 24–72 h, representing the seven major zones (100, 010, 001, 111, 110, 101, 011). The data corresponding to the other zones is more difficult to collect and interpret, as the tropomyosin filaments only diffract strongly in a few directions. The film data were digitized on an Optronics P1000 scanner at a resolution of 100 μm. At this resolution, the pixel-to-pixel separation represented 1/4,800 $Å^{-1}$ of reciprocal space.

The diffraction pattern from a crystal is proportional to the square of the Fourier transform of the unit cell structure, averaged over all the unit cells irradiated by the x-ray beam, and over the time frame of the photograph. This averaging is reproduced in the simulations by Fourier transforming a large array containing a number of unit cells. A unit cell is represented by an array containing the electron density in the unit cell, sampled at regular intervals. "Motion" in the unit cell is included by modifying the electron density array to represent the molecules in the cell displaced from their average positions, frozen at some instant in time. A superlattice of such arrays, containing a distribution of molecules in various frozen states of motion, represents the crystal. The distribution of states can be adjusted so that the root mean square displacement of the atoms is some appropriate value. Displacements correlated over more than one unit cell were represented by frozen "waves" of displacements running through the crystal along the filament directions. The justification for using waves of displacements stems from analyses of small molecule crystals, where the existence of phonons have been conclusively demonstrated.

In the simulations performed here, the unit cell was represented by a 16 × 32 × 32 array, and the final superlattice was 256 × 512 × 512. The array sizes were chosen so that there was sufficient sampling between Bragg spots in the computed diffraction pattern, and the data extended out to 15 Å resolution. The superlattice array was Fourier-transformed, and appropriate sections of it were extracted to match the experimental data views, as described below.

The second set of simulations were performed using a different approach. The Patterson function, or autocorrelation function of the atomic position vectors, is well resolved in an ideal crystal, but becomes less well defined away from the origin in a disordered crystal. The loss of sharpness in the function is related to the amplitude and correlations of the disorder. In the simulations performed here, the Patterson function for an ideal crystal is multiplied by a truncation function to simulate the disorder caused by the filament motions, similar to the technique used by Caspar et al. (3) to simulate the diffuse scattering from insulin. The diffuse intensities for the disordered crystal are calculated using the expression:

$$I_D = (1 - e^{-B\sin^2\theta/\lambda^2})\, FT[P_0\Gamma(r)], \qquad (1)$$

where P_0 is the Patterson function for the ideal crystal and $\Gamma(r)$ is the truncation function. The form of the truncation function provides some information about the nature of the correlations, and in agreement with Caspar et al., we found that a Gaussian truncation of the form $\Gamma(r) = e^{-r/\gamma}$ gave the best fit to the observed diffuse scattering. Thus, the probability of the motion of two atoms being correlated is related by an inverse exponent to the distance between them.

The motion in tropomyosin crystals is strongly anisotropic and inhomogenous. The filament axes of the symmetry-related molecules in the unit cell run in different directions, and its motion is related to this direction, so that no simple function can be used to describe all the motions in the unit cell. An earlier set of simulations made the assumption that the diffraction from a single molecule appeared primarily in two octants of reciprocal space, and this assumption was used to separate the diffraction from the different molecules. Thus we were able to treat the motions of each molecule separately. The simulated data was qualitatively similar to the experimental data, but the assumption used is not strictly valid, as the diffraction from each molecule appears to some degree in all the octants of reciprocal space.

In the simulations presented here, the Patterson function for a single molecule in the unit cell was truncated according to a model for the motion of that molecule. It was then Fourier-transformed to obtain the diffraction intensities, and symmetry-averaged over all the octants to produce the contributions of the other three symmetry-related molecules. This approach has the advantage that one is dealing with the motions of a single molecule and it is intuitively easier to design and interpret the appropriate truncation function. However, there are expected differences between the observed and simulated diffraction patterns that are due to the lack of interference effects between the molecules.

The final truncation function, applied to the ideal Patterson function of a single molecule, consists of two terms, one dealing with the short-range and one with the long-range correlated motions. For the short-range correlated motions, the Patterson was truncated by an expression of the form $\Gamma_s = e^{-\sqrt{(d/\gamma_{ds})^2 + (f/\gamma_{fs})^2}}$, where d is the distance along the filament and f is the distance perpendicular to the filament direction. γ_{ds} and γ_{fs} are the Gaussian truncation parameters for the short-range correlated motions along and perpendicular to the filament direction respectively. If γ_{fs} and γ_{ds} are set equal, the expression represents a spherical truncation, or isotropic, homogenous disorder.

For the long-range correlated motions that give rise to the Bragg-coupled diffuse scattering, the situation is more complex. A variety of different truncation directions and parameters were tested, corresponding to different models for the long-range correlated motions. The most successful simulations were those in which the correlations were along the axes of the cross-connecting filaments in the lattice. There are four molecules in the asymmetric unit, and therefore three molecular axes in addition to the molecule that is undergoing the displacements. Thus, this part of the truncation function consists of three terms, each truncating the Patterson function by an expression of the same form as the short-range correlations, but with directions along different lattice connections. In this case, however, d is the distance along the axis of the cross-connecting filament, and f the distance perpendicular to this axis. Thus the motion of any atom in a

filament is correlated to the motions of other atoms that lie along the directions of the three cross-connecting filaments.

The final form of the truncation function is a combination of the two terms, so that

$$\Gamma(r) = p \times \Gamma_s + (1 - p) \times \Gamma_l,$$

where p is the proportion of short-range correlated motion, and $(1 - p)$, correspondingly, the proportion of long-range correlated motion. p is a variable which is also refined during the simulations, and the two Γ terms represent the short and long range correlations respectively. The raw output from both the superlattice and the analytical simulations consists of a 256 × 512 × 512 array which contains a part of reciprocal space sampled at intervals of $1/1{,}920$ Å^{-1} intervals in the a^* direction, $1/3{,}840$ Å^{-1} in the b^* direction, and $1/4{,}768$ Å^{-1} in the c^* direction. Sections of this array corresponding to the experimental data were obtained by reverse-ray-tracing from the simulated film to find the reciprocal lattice coordinates that were intersected by Ewald's sphere for the chosen crystal orientation. These coordinates do not usually lie at the reciprocal lattice sample points, and so the diffraction intensity was found by interpolating between the 64 nearest sampled positions, using a Gaussian interpolation routine.

The initial comparisons of the experimental and calculated data were purely qualitative, and there was some difficulty in finding an appropriate quantitative comparison technique. An 'R-factor' which compared the two data sets in a pixel-to-pixel fashion was found to be too sensitive to artifacts in the experimental data, as was a correlation coefficient. The best method we found for the comparison was to block out the Bragg regions of each pair of data pictures and subtract the remaining simulation data from the experimental data (see Fig. 2). This 'difference' map was found to provide an easily interpretable indication of the differences between the two sets of data, and it was possible to estimate the simulation parameters that gave the best fit, although it is not a truly quantitative measure. The parameters for the analytical simulations were also derived by a visual matching of the observed and calculated patterns.

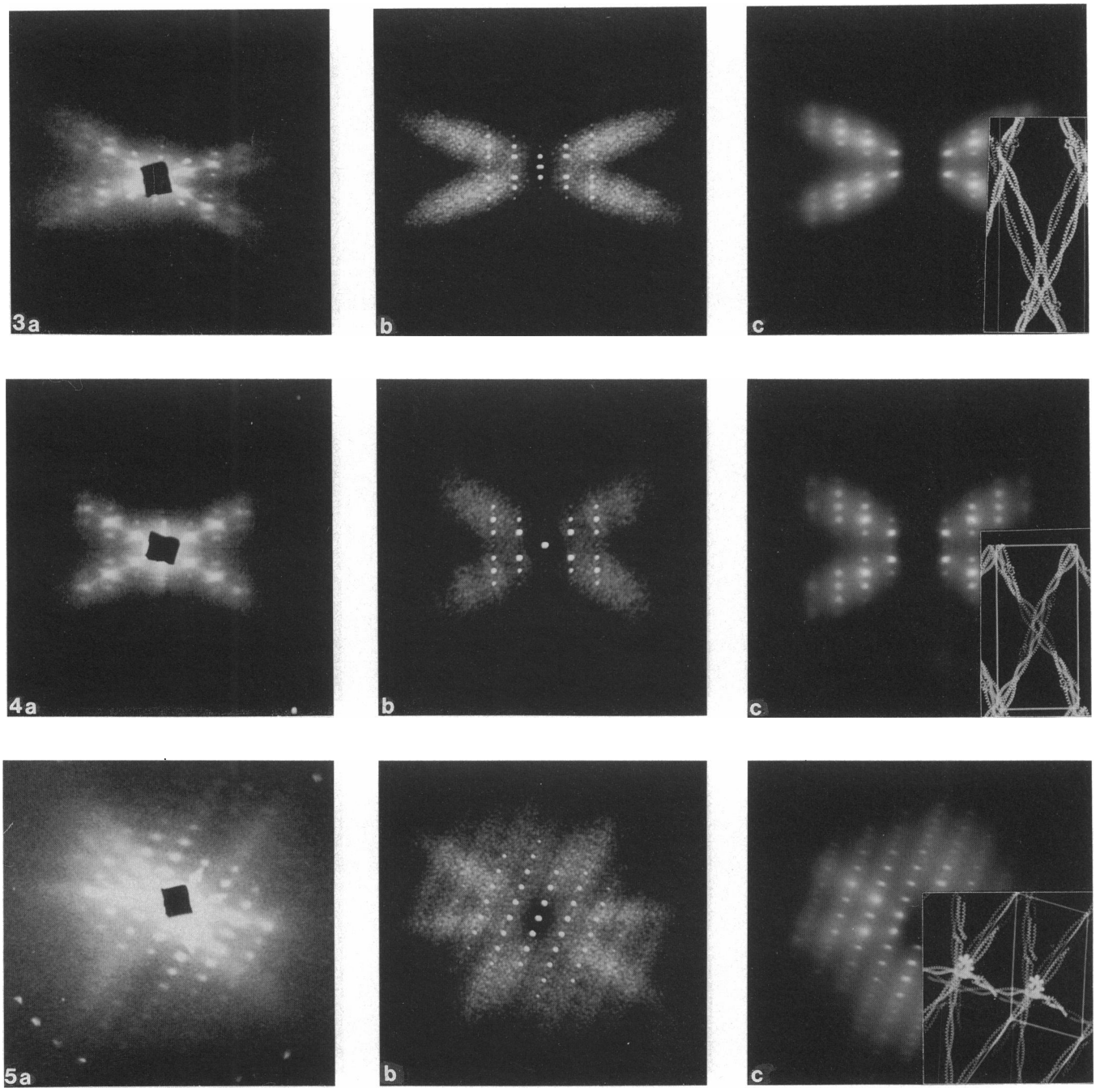
RESULTS AND DISCUSSION

An intuitive understanding of the three-dimensional diffuse scattering patterns can be obtained by examining the data and the filament directions in each of the views corresponding to the experimentally measured ones (see Figs. 1, 3–8). These structural projections are obtained by projecting the known three-dimensional structure into a plane perpendicular to the x-ray beam direction in the corresponding experimental data photograph. Thus, the 100 data photograph corresponds to a projection of the structure in the $b^* c^*$ plane, or perpendicular to the a axis. The insets in Figs. 1 *c*, 3 *c*–8 *c* show the appropriate crystal structure projections corresponding to the data photographs in Figs. 1 *a*, 3 *a*–8 *a*.

It is clear that the diffuse patterns in the different views are quite distinct from one another. However, there are some basic features in common to all the views. All the data photographs, to a greater or lesser extent, show two discrete kinds of diffuse scattering; the large clouds spread all over reciprocal space, and the smaller streaks associated with the Bragg spots. The size of diffuse scattering haloes is inversely related to the distance over which the disorder is coupled in the crystal; for example, disorder that was coupled over three unit cells would produce small haloes around the Bragg spots; the size of the halo being the inverse of three unit cell lengths. Disorder coupled over a fraction of the unit cell would produce large haloes around the Bragg spots which would overlap with the haloes from nearby spots, thus producing the clouds of diffuse scattering that are not associated with Bragg spots. The two kinds of diffuse scattering present in tropomyosin then indicate that there are two discrete ranges of correlation distances. The Bragg-coupled streaks are about the size of the Bragg spot separation, so that one range of correlation distances is of the order of the unit cell dimensions or greater. The other range is considerably less than the unit cell dimension, and produces the larger diffuse clouds.

A closer examination of the data photographs can provide still more information about the motion in the crystal. The diffuse scattering is closely related to the molecular transform, and so both the Bragg and diffuse scattering associated with a particular part of the unit cell will appear in the same region of reciprocal space. In the 100 view, therefore, the symmetry in the four quadrants of the data picture correspond to the symmetry of the molecules in the structure. One pair of molecules runs along each diagonal of the rectangular plane, and produces the diffuse scattering along one diagonal of the data photograph. The long and short arms of each molecule run in slightly different directions, and produce the two slightly divergent clouds seen in each quadrant. The fact that the diffuse scattering is symmetrical means, not surprisingly, that the symmetry of the displacements is the same as the symmetry of the molecular packing.

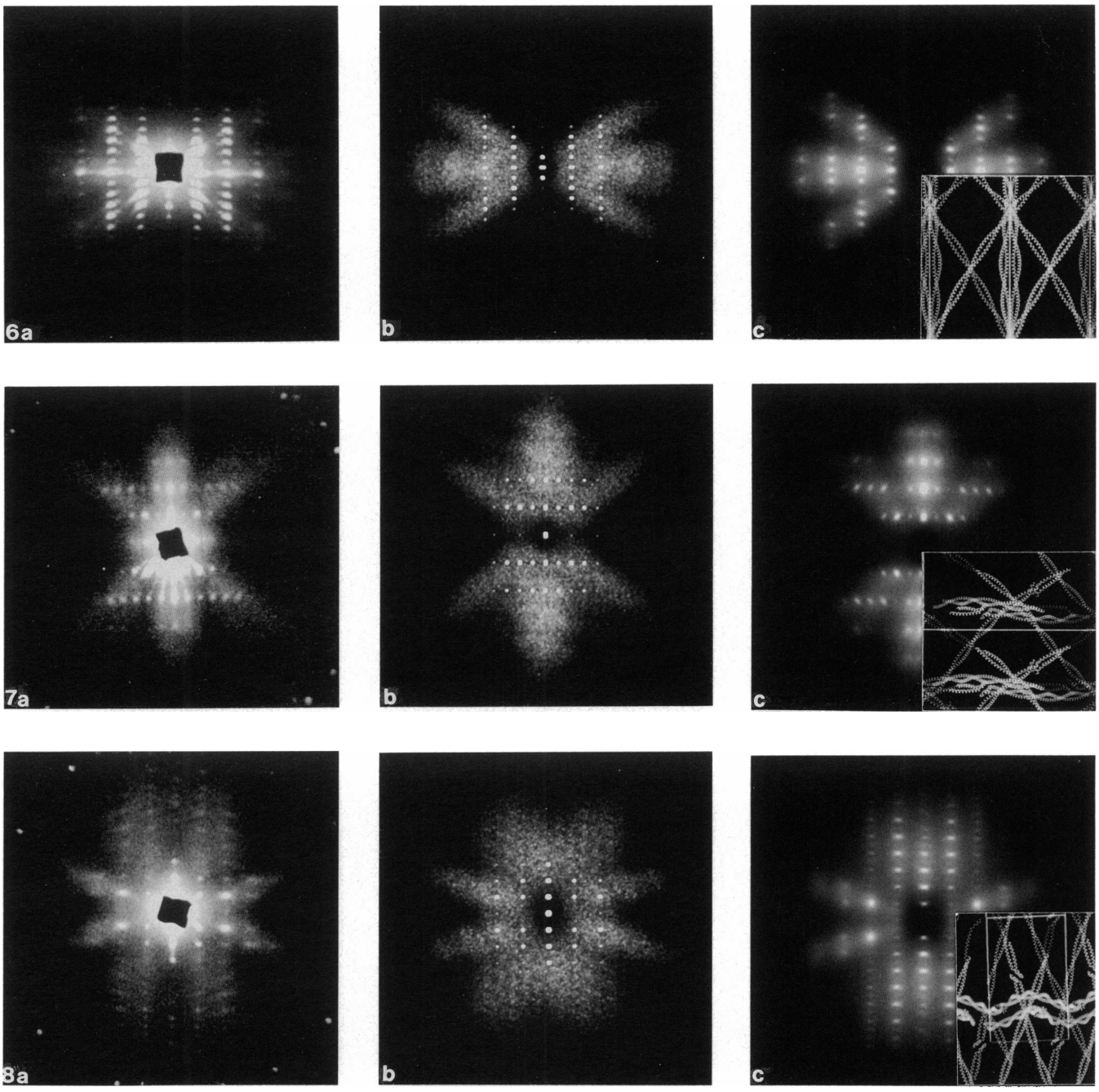
Further information can be deduced from examining the Bragg-coupled streaks in this photograph. These streaks, as mentioned before, are due to long-range correlated motions of the molecules. Isotropic motions would produce isotropic haloes about the Bragg spots, which is clearly not the case (see Fig. 1 *a*). The marked asymmetry of the streaks implies a corresponding anisotropy of the motions. Motions that are propagated in a crystal produce diffuse scattering that is elongated in the direction of propagation of the motion, whether it is propagated as a longitudinal or transverse wave. The streaks are thus consistent with either a longitudinal or transverse wave running along the filament directions. The force required to bend a long helical molecule such as tropomyosin is clearly less than that required to stretch or compress it (20) so that longitudinal waves are less easily propagated than transverse waves running along the filament. Theoretical studies of wave propagation in crystals of organic acids (21), which have chain



FIGURES 3, 4, and 5 Diffraction corresponding to 010, 001, and 111 views respectively. 3 *a*, 4 *a*, 5 *a*: experimental data. 3 *b*, 4 *b*, 5 *b*: simulations using superlattice technique. 3 *c*, 4 *c*, 5 *c*: simulations using analytical technique. (*Insets*) Corresponding projections of the molecular structure.

structures, have also shown that longitudinal waves running along the chains are virtually impossible. Transverse waves would give rise to discs of diffuse scattering around the direction of diffraction of the filaments. At present, our experimental data are too incomplete to allow us to observe these discs, but the sections that we have collected are consistent with their presence. It may be deduced, then, that the observed diffuse scattering streaks are consistent with transverse waves of displacements running along the filament directions. A similar intuitive explanation can be made for the shapes of the diffuse scattering in each of the other views.

This qualitative approach to the analysis of the diffuse scattering can provide important clues as to which parts



FIGURES 6, 7, and 8 Diffraction corresponding to 110, 101, and 011 views respectively. 6 *a*, 7 *a*, 8 *a*: experimental data. 6 *b*, 7 *b*, 8 *b*: simulations using superlattice technique. 6 *c*, 7 *c*, 8 *c*: simulations using analytical technique. (*Insets*) Corresponding projections of the molecular structure.

of the unit cell are moving, and in which direction. This type of analysis would be less easily interpretable for a globular protein than for tropomyosin, where the diffraction from different molecules and different parts of the molecule is conveniently separated in reciprocal space. The shape of the diffuse scattering gives an indication of the anisotropy of the motions.

Comparison of the techniques

The superlattice approach was very successful in simulating the short-range correlated motions, but was not able to simulate the long-range correlated motions very well. Conversely, the analytical method was able to provide a good match for the Bragg-coupled diffuse streaks, but

not as good a match for the more cloudy diffuse scattering. The explanation for this lies in the different models of the dynamic structure of tropomyosin that were used in the two techniques, and the strengths and limitations of each technique.

In the superlattice method, the molecular model was that of a vibrating string, fixed at the node points where the molecule comes close to the other molecules. The motion was thus, entirely in the long and short arm, and the node regions stayed stationary. This is a reasonable approximation for motions correlated over short distances, where the more constrained node regions can be imagined to act as damping regions for propagation of the motions. In the center of the long arm, the molecule is about 150 Å away from any other molecule, whereas in the node regions, the molecule are separated by about the thickness of the filament, i.e., 20 Å. Clearly, the motions will have greater amplitude in the arm regions than in the node regions, and the B-factors calculated from the Bragg data bear this out. Hence the assumption of fixed nodes is reasonable for the short-range correlated motions.

The long-range correlated motions are correlated over distances of a few hundred Å, as evidenced by the size of the Bragg-coupled streaks and the simulation parameters. These motions must necessarily include some correlations through the node regions as well, since there are two node regions along every molecular length. Hence, a model which assumes immovable nodes is less reasonable for long-range correlated motions. Also, the procedure for modeling long-range correlated motions by this method limits the correlation distance to an integral number of unit cell lengths. In view of these problems, it is not surprising that the superlattice technique failed to provide a good match for the Bragg-coupled streaks in the diffuse data.

The model for the analytical method assumes homogeneity of the motions along the molecule. That is, all parts of the molecule have the same amplitude of motion perpendicular to the molecular axis, and the correlations are uniform throughout the molecular length, so that any part of the molecule is correlated to all the other parts up to the correlation distance. The node regions also encompass and take part in the motion. This is a more reasonable model for the long-range correlated motions, and explains the good match of the Bragg-coupled streaks in the simulations and experimental data. In addition, this method is particularly suitable for fine-tuning the correlation distances and modeling correlations in any direction.

The B-factor data shows that at least some part of the motion must have smaller amplitude at the node regions than in the arm regions. On the basis of the simulations of Bragg-coupled streaks, we can assume that to a first approximation the long-range correlated motions have uniform amplitude along the molecular length, and so the short-range correlated motions must have considerably smaller amplitudes at the node regions. However, the analytical method is at present not able to include inhomogeneity in the model for molecular motions, and so the simulations of short-range correlated motions assumed uniform amplitudes of motion along the molecule. The approximate position and shape of the diffuse cloud was correct, but the details of the diffuse intensity were not well matched, and this can be explained by the shortcomings of the model used.

How can we improve the model for the simulations? The superlattice simulations are based upon a library of molecules which have displaced paths according to the vibrating string model. More sophisticated models might include smaller displacement amplitudes at the nodes, so that the motions were effectively damped at these regions. Molecular dynamics packages could be used to generate physically reasonable displaced molecular paths, which are not subject to constraints as severe as the present models.

The B-factors calculated for tropomyosin indicated that there is significant anisotropy in the plane perpendicular to the molecular axis, i.e., that the molecular motion in this plane was not uniform in all directions. The superlattice simulations were not able to distinguish this anisotropy, but it is undoubtedly represented in some form in the diffuse scattering. More sophisticated comparison techniques may allow us to make deductions about the anisotropy as well.

The analytical procedure is easily modified to include anisotropy. The models used here assumed isotropic motion in the plane perpendicular to the molecular axis, but modifying this is fairly trivial, and may provide useful information in the future. The more serious problem with the analytical method is the homogeneity implicit in the models. The simulations of tropomyosin tackled the problem of the intermolecular inhomogeneity by working with the Patterson function from a single molecule. This effectively reduced the problem to the dynamics of a single molecule, the dynamics of the other molecules being deduced by symmetry. However, the question of intramolecular inhomogeneity has not been solved. One might be tempted to divide the problem further, in analogy with the intermolecular case, and assume that the diffraction from the node regions was independent of that from the two arm regions. However, the constructive and destructive interference from all these regions contributes to the diffraction pattern, and so the calculated Bragg as well as diffuse scattering would be

TABLE 1 **Correlation distances (CDs) determined from the simulations**

Technique	Short-range correlated		Long-range correlated	
	∥ filament axis	⊥ filament axis	∥ cross-connecting filament axes	⊥ cross-connecting filament axes
Superlattice	63 Å	25 Å		
Analytical	65 Å	35 Å	125 Å	450 Å

The superlattice simulations were not successful in simulating the Bragg-coupled streaks; thus, no values are reported for the long-range correlation distances. The CDs reported in the table for the analytical simulations are directly from the simulation parameters. The CDs reported for the superlattice simulations are estimates, because this technique does not directly provide these parameters. For the superlattice simulations, therefore, $CD_S^{\perp}$ is estimated as the thickness of the filament. $CD_S^{\parallel}$ is an average of the correlation distances over the molecular length.

incorrect. Hence, a more sophisticated approach must be devised to take care of this problem.

In summary, neither technique is suitable for simulating the entire diffuse scattering data at present, but each explains a key element of the diffuse scattering. In fact, our combination of superlattice and analytical techniques provided the parameters describing the proportions of long and short-range correlated motions, the correlation distances, and the amplitudes of motion.

The long and short-range motions are in the proportions 0.05:0.95. This estimate agrees with that obtained from similar diffuse scattering studies on insulin and lysozyme (3, 16), as well as the Mössbauer studies on myoglobin (22). Thus, in proteins studied so far, most of the displacement of a given atom is due to the short-range correlated motions.

Correlation distances

Correlation distances are most easily obtained from the analytical method, which permits fine-tuning of the parameters more easily than the superlattice method. The correlation distances for the superlattice method were included into the basic model used, and modifying them was a non-trivial procedure. The 'best-fit' correlation distances obtained are listed in Table 1.

The symbol CD is used to represent the correlation distance in the discussion that follows. The superscripts $CD^{\parallel}$ and $CD^{\perp}$ refer to the CD parallel or perpendicular to the molecular axes, while the subscripts CD_S and CD_L refer to the short and long-range correlation distances. For the long-range correlated motions, the additional superscript CD^{cc} is used as a reminder that these motions are correlated mostly along the direction of the cross-connecting filaments in the lattice.

The parameters obtained for the short-range correlated motions from the two techniques are quite comparable (see Table 1). The correlation distance along the molecular axis (65 Å) is less than the length of the short-arm region but one must remember that this is an average over the whole length of the molecule, and that the node regions could have significantly smaller amplitudes. The value of $CD_S^{\parallel}$ from the two techniques is thus consistent with the hypothesis that the long arm and short arm motions are correlated within each region, but are independent of each other, and that the node regions are relatively unaffected by the short-range correlated motions. This is in agreement with earlier simulations of the tropomyosin diffuse scattering by Boylan and Phillips (14).

The correlation distances perpendicular to the molecular axis are harder to interpret. $CD_S^{\parallel}$ from the analytical technique is ~10 Å larger than the thickness of the molecule. It is tempting to hypothesize that this figure represents the molecular thickness plus a 5 Å thick 'shell' of water, which is being dragged along with the molecule as it moves. Alternatively, this could be some unidentified artifact of the simulation technique.

The parameters for the long-range correlated motions were obtained only from the analytical simulations. The main correlation direction was along the axes of each of the three cross-connecting molecules, and the correlation distance was about 450 Å, slightly more than a molecular length. Thus we can visualize transverse waves of displacements running along the filaments, the displacements of these filaments being correlated along the directions in which the cross-connecting molecules run. The node regions, where the cross-connecting molecules come close to one another, must obviously be instrumental in transferring the displacement effects from one molecule to another.

Amplitudes

The average amplitudes of the molecular displacements are easily obtained from the superlattice simulations, where the superlattice consists of a distribution of molecules with displaced paths. The best simulations were determined on the basis of the difference maps, and the amplitudes obtained are listed in Table 2.

The amplitudes of the motions are incorporated into the analytical method in the term $(1 - e^{-B\sin^2\theta/\lambda^2})$ (see Eq. 1), where the B-factor is proportional to the square of the average displacement. In theory, one could refine the B-factor value from the diffuse scattering data just as the correlation distances were refined, and thus calculate the amplitudes of motion. In practice, however, it is

TABLE 2 **Average amplitudes of motion calculated from Bragg data and two kinds of simulations**

		Superlattice simulations	Analytical simulations	
	Bragg data (B-factors)	short-range correlated	long-range correlated	short-range correlated
Average	5.4 Å	3.1 Å	1.2 Å	5.2 Å
Long arm average	5.8 Å	3.5 Å		
Short arm average	4.9 Å	2.5 Å		
Long arm peak	7.0 Å	6.5 Å		
Short arm peak	6.2 Å	4.4 Å		

Blank spaces in the table indicate that the appropriate values could not be calculated from the analytical simulations.

not easy to distinguish the change in the calculated diffraction patterns as B is changed. The B value used in the simulations was 2,000 Å^2, which is consistent with the average B calculated from the Bragg diffraction data, but this study cannot be considered as a totally independent determination of B.

Although the analytical method does not provide any amplitude information per se, it can be deduced using some assumptions. One assumption is that all the displacements are due to the short and long-range correlated motions, and hence that the B-factor data represent the sum of the average amplitudes of motion of these motions. To a first approximation, we get

$$\frac{I_{D_L}}{I_{D_S}} = \frac{B_L}{B_S} = \frac{x_L^2}{x_S^2} \tag{2}$$

$$B = 8\pi^2 x^2 = 8\pi^2(x_L^2 + x_S^2), \tag{3}$$

where I_{D_L} represents the diffuse intensity due to the long-range correlated motions, B_L is the B-factor for these motions, and x_L is the displacement due to these motions. I_{D_S}, B_S and x_S are the corresponding terms for the short-range correlated motions.

Using the above relations, we can estimate the amplitudes of motion for the short and long range correlated motions. The results of the simulations indicated that the relative proportions of the motions was 0.95::0.05, i.e. that 95% of the diffuse intensity was associated with short-range correlated motions. Also, the average of the B-factors, B, calculated from Bragg diffraction data (23) is 2,278, corresponding to an average movement, x, of 5.4 Å. Hence B_S is 2,164 Å^2 and B_L is 114 Å^2. The average amplitudes of motion are thus x_L = 1.2 Å and x_S = 5.2 Å.

The information about average atomic displacements is condensed into Table 2. The table shows that the results from the superlattice simulations are consistently about 2 Å less than those from the Bragg data. One might assume that this is because only part of the total motion (the short-range correlated part) is simulated by this technique. However, the analytical simulations inform us that only 5% of the total motion is associated with long-range correlations, and so this cannot explain the discrepancy. It must be remembered that the 'best' parameters for the superlattice simulations were determined on the basis of the 'difference maps', and are accurate only to ~1 Å. Another source of this difference could be the inaccuracy of the Bragg B-factors at less than atomic resolution.

The need for more sophisticated comparison techniques is evident. Ideally, both Bragg and diffuse data, on the same absolute scale, would be included in the comparisons. If the simulated Bragg data were scaled to the experimental data, the corresponding diffuse intensities could be compared, providing a more precise determination of the best simulation parameters. However, the Bragg data is necessarily overexposed in order to record the diffuse data on film, and so with film recording it is practically impossible to obtain accurate diffuse as well as Bragg data. Some hope is offered by imaging plate technology, which appears to have the dynamic range required to record both Bragg and diffuse data on the same frame and also have no background fog as is seen on film. In fact, imaging plates have been used to record diffuse data from yeast initiator tRNA crystals (12) and in the future it may be possible to obtain higher quality tropomyosin data using imaging plates.

What do the parameters from these analyses mean? In the case of the short range correlations, they report the size of the part of the protein that is moving as a more-or-less rigid body. The correlation distance can be used to discriminate between independent motions of atoms (no correlation), liquidlike motions (correlations of the order of a few Å), or the movement of entire domains (intermediate correlation lengths). These are often biologically important distinctions, and can be determined by diffuse scattering analyses. The longer range correlations are lattice-dependent phenomena, and are less interesting to the biologist. However, they are important to the crystallographer, who is trying to obtain the best quantitative interpretation of the x-ray diffraction data.

In summary, the tropomyosin molecules are considerably more mobile than most globular protein molecules, which usually have amplitudes of the order of 1 Å or less. The motions of tropomyosin are marked by a high degree of both anisotropy and inhomogeneity, but all the molecules are apparently identical in their amplitudes and modes of motion. Each filament has two types of motion. One has large amplitudes and is composed of transverse motions of the long and short arms. The motions of each of the arms is independent of the other. The node regions have very little motion and can be

thought of as points where the filament string is pinned down. In view of the large ($\simeq$5–7 Å) amplitudes of the arm motions and the relatively small distance between molecules in the node regions ($\simeq$20–25 Å), it seems unlikely that the molecules in the node regions were involved in these large scale motions. The fluctuations may be anisotropic in the plane perpendicular to the filament axis, but this could not be determined from the diffuse scattering analysis. The second type of motion has much smaller amplitudes, and the motions are correlated over a little more than a molecular length. These small fluctuations propagate through the nodes as well as the arm regions, and the correlations extend in the directions of the cross-connecting molecules in the lattice.

This diffuse scattering analysis has provided a quantitative understanding of the correlated molecular motions within the tropomyosin crystal, and has allowed us to make some statements about the intrinsic flexibility of the tropomyosin molecule. Furthermore, the methods of analysis of diffuse scattering described here may be useful for characterizing the dynamics of other proteins in crystals.

We thank Don Caspar and Carolyn Cohen for inspiration and helpful discussions.

This work was supported by grants from National Institutes of Health (AR32764), National Science Foundation (DMB 87-16507), the Robert A. Welch Foundation (C-1142) and the W. M. Keck Foundation. S. Chacko was supported by a Predoctoral Fellowship from the W. M. Keck Center for Computational Biology.

Received for publication 29 July 1991 and in final form 9 November 1991.

REFERENCES

1. Frauenfelder, H., G. A. Petsko, and D. Tsernoglou. 1979. Temperature-dependent x-ray diffraction as a probe of protein structural dynamics. *Nature* (*Lond.*). 280:558–563.
2. Hartmann, H., F. Parak, W. Steigemann, G. A. Petsko, D. Ringe Ponzi, and H. Frauenfelder. 1982. Conformational substates in a protein: structure and dynamics of metmyoglobin at 80K. *Proc. Natl. Acad. Sci. USA.* 79:4967–4971.
3. Caspar, D. L. D., J. Clarage, D. M. Salunke, and M. Clarage. 1988. Liquid-like movements in crystalline insulin. *Nature* (*Lond.*). 332:659–662.
4. Phillips, G. N., Jr., J. P. Fillers, and C. Cohen. 1980. Motions of tropomyosin. *Biophys. J.* 10:485–502.
5. Petsko, G. A., and D. Ringe. 1984. Fluctuations in protein structure from x-ray diffraction. *Annu. Rev. Biophys. Biochem.* 13:331–371.
6. Debye, P. 1914. *Ann. d. Physik.* 39:789. *Translated in* The Collected Papers of Peter J. W. Debye, Interscience Publishers, New York. 1954.
7. Einstein, A. 1907. Planck's theory of radiation and the theory of specific heat. *Ann. d. Physik.* 22:180–190. *Translated in* The Collected Papers of Albert Einstein, Vol. 2. Princeton University Press, New Jersey.
8. James, R. W. 1982. *In* The Optical Principles of the Diffraction of X-rays. Oxbow Press, Connecticut. 193–230.
9. Doucet, J., and J.-P. Benoit. 1987. Molecular dynamics studied by analysis of the x-ray diffuse scattering from lysozyme crystals. *Nature* (*Lond.*). 325:643–646.
10. Clarage, J. B., M. S. Clarage, W. C. Phillips, R. M. Sweet, and D. L. D. Caspar. 1992. Correlations of atomic movements in lysozyme crystals. *Proteins.* 12:145–157.
11. Helliwell, J. R., I. D. Glover, A. Jones, E. Pantos, and D. S. Moss. 1986. Protein dynamics: use of computer graphics and protein crystal diffuse scattering recorded with synchrotron x-radiation. *Biochem. Soc. Trans.* 14:653–655.
12. Kolatkar, A., G. N. Phillips, Jr., R. Basavappa, and P. B. Sigler. 1990. Diffuse scattering analysis of crystalline yeast initiator t-RNA. Meeting Proceedings of the American Crystallographic Association. 18:94.
13. Doucet, J., J.-P. Benoit, W. T. B. Cruse, T. Prange, and O. Kennard. 1989. Co-existence of A- and B-form DNA in a single crystal lattice. *Nature* (*Lond.*). 337:190–192.
14. Boylan, D., and G. N. Phillips. 1986. Motions of tropomyosin. *Biophys. J.* 49:76–78.
15. Phillips, G. N., Jr., E. E. Lattman, P. Cummins, K. Y. Lee, and C. Cohen. 1979. Crystal structure and molecular interactions of tropomyosin. *Nature* (*Lond.*). 278:413–417.
16. Clarage, J. 1990. Disorder in protein crystals. Ph.D. thesis. Brandeis University, Waltham, MA.
17. White, S. P. 1987. Structure of co-crystals of tropomyosin and troponin. Ph.D. thesis. University of Illinois, Urbana-Champaign, IL.
18. Bailey, K. 1948. *Biochem. J.* 43:271–272.
19. Phillips, G. N. 1985. *Methods Enzymol.* 114A:128–131.
20. Levy, R. M., and M. Karplus. 1979. Vibrational approach to the dynamics of an α-helix. *Biopolymers.* 18:2465–2495.
21. Amorós, J. L., and M. Amorós. 1968. In Molecular Crystals: Their Transforms and Diffuse Scattering. John Wiley & Sons, Inc., New York.
22. Nienhaus, G. U., J. Heinz, E. Huenges, and F. Parak. 1989. Protein crystal dynamics studies by time-resolved analysis of x-ray diffuse scattering. *Nature* (*Lond.*). 338:665–666.
23. Phillips, G. N., Jr., J. P. Fillers, and C. Cohen. 1986. Tropomyosin crystal structure and muscle regulation. *J. Mol. Biol.* 192:111–131.