

Quantitation of Molecular Densities by Cryo-electron Microscopy

Determination of the Radial Density Distribution of Tobacco Mosaic Virus

Michael F. Smith and John P. Langmore†

*Biophysics Research Division and Department of Biological Sciences, The University of Michigan,
2200 Bonisteel Blvd, Ann Arbor, MI 48109-2099, U.S.A.*

(Received 12 October 1991; accepted 31 March 1992)

We have determined the absolute mass and radial scattering density distribution of tobacco mosaic virus in the frozen-hydrated state by energy-filtered low-dose bright-field transmission electron microscopy. The absolute magnitude of electron scattering from tobacco mosaic virus in 150 nm of ice was within 3.0% of that predicted, with inelastic scattering accounting for ~80% of the scattering contrast. In order to test the accuracy of the radial reconstruction, a computer model of tobacco mosaic virus was built from the atomic co-ordinates assuming uniform solvent density. The validity of the model was confirmed by comparison of X-ray scattering and predictions of the model (R factor = 0.05). First-order corrections for the microscope contrast transfer function were necessary and sufficient for conversion of the cryo-electron microscopy images into accurate representations of the mass density. At 1.9 nm resolution the compensated reconstruction and model had density peaks of similar magnitude at 2.4, 4.2, 6.0 and 7.8 nm radius and a central hole of 2 nm radius. Equatorial Fourier transforms of the corrected electron images were in excellent agreement with predictions of the model (R factor = 0.12). Thus, the uniform solvent approximation was adequate at 1.9 nm resolution to describe quantitatively X-ray scattering in liquid water and electron imaging in vitreous ice. This is the first demonstration that cryo-electron microscopy images can be used to quantitate the absolute mass, mass per unit length and internal density distributions of proteins and nucleic acids.

Keywords: inelastic scattering; energy-filtration; mass analysis; contrast transfer function; computer modeling

1. Introduction

The purpose of “image reconstruction” is to determine the mass densities within molecules by analysis of the intensities within images. The widespread application of electron microscopy (EM†) to represent the internal structure of biological molecules makes it important to evaluate the accuracy of those representations.

Most specimens used for EM are stained with heavy atoms to enhance image contrast and preserve the structure against dehydration and radiation damage. Although the images represent

the sizes, shapes and locations of biological molecules, they are not useful for quantitation of mass, mass per unit length or mass per unit volume, because the image intensities are dominated by scattering from stain bound to or excluded from the molecules.

Bright-field and dark-field images of unstained dehydrated molecules have been used for quantitative microscopy (e.g. see Zeitler & Bahr, 1957, 1962; Brakenhoff, 1974; Lamvik & Langmore, 1977). The best information can be obtained from images recorded by dark-field scanning transmission electron microscopy (STEM) of unstained, frozen-dried biological specimens (for a review, see Wall & Hainfeld, 1986). STEM analyses have the advantages of high image contrast without the use of heavy-atom stains or phase contrast, and direct electronic recording of the images. The disadvantage of STEM is the distortion caused by

†Author to whom all correspondence should be addressed.

‡Abbreviations used: EM, electron microscopy; STEM, scanning transmission EM; CTF, contrast transfer function; TMV, tobacco mosaic virus.

freeze-drying, which limits resolution to 2 to 4 nm (Wall & Hainfeld, 1986).

Methods to preserve and image biological molecules in thin layers of vitreous ice have created new possibilities for quantitative EM. Vitrification of biological specimens minimizes many of the artifacts associated with other preparative techniques (Taylor & Glaeser, 1974, 1976; Dubochet *et al.*, 1982; for reviews, see Chiu, 1986; Stewart & Vigers, 1986; Dubochet *et al.*, 1988). Bright-field imaging of frozen-hydrated specimens has emerged as a reliable, high-resolution technique to examine molecules in their native hydrated conformation. Cryo-EM images, however, have low contrast because of the similarity between the density of biological material and the density of vitreous ice. The images also contain a high background of inelastic scattering. To overcome the low signal and high background, cryo-images are typically recorded 1 to 3 μm underfocus to maximize phase contrast. The introduction of phase contrast, however, significantly complicates the quantitative analysis of the images.

The phase contrast in images of thin specimens can be predicted by the first-order theory of image formation (Erickson, 1971; Erickson & Klug, 1971). However, the theory neglects inelastic scattering, which ordinarily contributes a majority of the scattered electrons in the images of frozen-hydrated samples (Langmore & Athey, 1987; Langmore *et al.*, 1990; Schröder *et al.*, 1990; Langmore & Smith, 1992). The primary effect of inelastic scattering is to reduce phase contrast by reducing the coherence of the transmitted electrons. A secondary effect is to increase scattering contrast by chromatic aberration in the focusing of the energy-loss electrons. Removal of the inelastically scattered electrons by use of an energy filter reduces image background (increasing contrast and reducing statistical noise) and provides beneficial inelastic scattering contrast (Langmore & Smith, 1992).

The relationship between object density, and phase and scattering contrast is commonly described by the microscope contrast transfer function (CTF). Under the defocused conditions that impart beneficial phase contrast to the images, the image intensities bear little resemblance to the object, due to the dependence of the CTF upon spatial frequency. Compensation (i.e. correction) for the CTF is theoretically possible, but implementation has been rare (Erickson & Klug, 1971; Toyoshima & Unwin, 1988, 1990). Usually the image intensities are compensated for phase contrast only, not compensated at all, or taken from electron or X-ray diffraction results. The first two approaches have been rationalized by arguments that the CTF has little effect at low resolution. Although the theoretical CTF for different imaging conditions has been illustrated (e.g. see Stewart & Vigers, 1986), the experimental effects of the CTF upon the images have not been studied except by Lepault & Leonard (1985), who showed that defocus had a substantial effect on the images of bacterio-

phage T4 tails and that for low spatial frequencies ignoring the CTF was better than compensating for phase contrast alone. The lack of confidence in compensation of the CTF is illustrated by the high-resolution cryo-EM study of tobacco mosaic virus (TMV) by Jeng *et al.* (1989). They concluded that "the amplitudes determined by electron microscopy are in error, because of the difficulties in correcting accurately for the CTF." Therefore, despite the fact that EM is thought to be most reliable at low resolution, Jeng *et al.* (1989) discarded image information at spatial frequencies less than 0.2 nm^{-1} and substituted information from X-ray crystallography, which is usually regarded as being inaccurate at low resolution. Despite numerous studies of biological molecules in ice, there has never been a quantitative evaluation of the compensated images of biological molecules of known structure.

We used TMV to verify the accuracy of using electron image intensities to determine the absolute internal densities of a frozen-hydrated biological molecule. A computer model of TMV was built from atomic co-ordinates assuming uniform solvent density. The validity of the model was confirmed by comparison of the predicted and experimental X-ray results at low resolution. A radial density distribution at 1.9 nm resolution was calculated from the compensated electron images and compared with the model. Compensation for the CTF was necessary and sufficient to accurately reconstruct molecular densities. The only empirical parameter necessary for compensation was determined from two images of the same molecules. The absolute values of contrast from RNA and protein, and the total amount of elastic and inelastic scattering agreed well with the predictions. The results confirm (at 1.9 nm resolution) the ability to model scattering of X-rays and the imaging of electrons using uniform solvent densities, and the ability of cryo-EM reconstructions to quantitatively represent the internal densities of proteins and nucleoproteins.

2. Materials and Methods

(a) Cryo-electron microscopy

TMV specimens were suspended over holey carbon films, frozen and transferred using standard techniques (Dubochet *et al.*, 1988). Bright-field images were recorded with a Zeiss EM902 operated at 80 keV using a 90 μm objective aperture. Kodak SO-163 was developed in full strength D-19 at 68°C for 12 min. The cold stage was maintained at 115 K during examination. Inelastically scattered electrons were filtered from the images by a spectrometer slit that had an effective width of 17 eV. Removal of inelastically scattered electrons increased image contrast about 4-fold, as reported (Langmore & Athey, 1987; Langmore *et al.*, 1990; Schröder *et al.*, 1990; Langmore & Smith, 1992). Specimen areas with vitreous ice were located at 3000 \times magnification. The microscope was focused at 85,000 \times magnification on an area $\sim 13 \mu\text{m}$ from the area of interest. Modifications to the microdose focusing system allowed for the pre-irradiation of the imaged area to be $< 3 \times 10^{-4}$ electrons/nm². Two low-dose images were made at 30,300 \times magnification with an

incident intensity of 200 electrons/nm² per s. The first image was recorded for 3 s at 780 nm underfocus, and was used for the radial density determination. The second image was recorded for 6 s at an additional underfocus of 3840 nm, and was used for the angle refinement and defocus calculations. The positions of the minima in the Fourier transform of the adjacent carbon film in the second image were used to determine the defocus of the images. Ice thickness was estimated to be 150 nm from the ratio of the ice background to the incident intensity, assuming mean free path lengths of 280 nm for elastic scattering and 180 nm for inelastic scattering (Langmore & Smith, 1992). The efficiency of removal of inelastic scattering from the images was 1.0 (Langmore & Smith, 1992).

(b) Image recording and processing

Electron micrographs were digitized as 512 × 386 pixel images using a STAR1 cooled CCD (Photometrics Ltd., Tucson, AZ, U.S.A.) equipped with a Nikon Micro-Nikor (focal length, 55 mm; aperture, f8.0). The CCD had contiguous 23 μm × 23 μm pixels. The micrographs were placed on a light box and digitized with a pixel size of 0.39 nm on the specimen. The density of the film image was calculated using the relationship $D_i(x,y) = -\log [T_i(x,y)/T_b(x,y)]$, where $T_i(x,y)$ is the transmitted light intensity on the film image and T_b is the transmitted light intensity at the same position on the light box without the image. The electron intensity was calculated using the relationship $I(x,y) = K [D_i(x,y) - D_f]$, where D_f is the density of the film without any electron exposure and K is the film speed. The value of K is about 0.45 electrons/μm². Magnification was calculated to be 30,300×, from the position of the 2.3 nm third layer line of TMV. Image processing was performed using the EMPRO program package (Smith *et al.*, 1990) on a Silicon Graphics IRIS 2500T workstation (Silicon Graphics, Inc., Mountain View, CA, U.S.A.).

(c) Computer modeling of TMV and solvent

TMV models were built on a Silicon Graphics IRIS 4D/220GTX from the atomic co-ordinates at 0.29 nm resolution (Namba *et al.*, 1989) obtained from the Brookhaven National Laboratory Protein Data Bank. The C-terminal residues 155 to 158 were not included in the co-ordinate map due to very high temperature factors (Namba *et al.*, 1989). Solvent scattering was modeled as described by Langridge *et al.* (1960), which assumes uniform solvent density and does not make any assumption about the exact boundary between solvent and biological molecule. The effective scattering from a particular amino acid or ribonucleotide was represented by the molecular scattering minus the scattering of an equal volume of solvent. The solvent density was assumed to be 1.0 g/cm³ for liquid water and 0.92 g/cm³ for vitreous ice. The solvent volume excluded by each amino acid was represented by the measured molecular volumes (Zamyatnin, 1972). The ribonucleotide volumes were calculated from the partial specific volume of RNA (Cammarano *et al.*, 1972) and the ribonucleotide molecular weight. Three-dimensional models were built by placing the amino acid and ribonucleotide effective molecular scattering at the co-ordinates of the α-carbons and phosphorus atoms, respectively. The results were the same if all non-hydrogen atom positions were used.

Electron scattering amplitudes were derived from the theoretical atomic scattering amplitudes in the forward

direction (from Schafer *et al.*, 1971). Electron scattering intensities were calculated from the elastic and inelastic scattering cross-sections. Elastic cross-sections were calculated from the scattering amplitudes reported by Schafer *et al.* (1971), assuming a 90 μm objective aperture. Inelastic cross-sections were estimated from the data of Wall *et al.* (1974), assuming a spectrometer efficiency of 100% as shown by Langmore & Smith (1992). Mass densities were derived by placing the amino acid and ribonucleotide molecular weights at the co-ordinates of the α-carbon and phosphorus atoms, respectively.

Two-dimensional models of the TMV scattering amplitudes, scattering intensities and mass densities were calculated by projection of the three-dimensional amplitudes, cross-sections and densities onto a plane. TMV projections were averaged over 6 different rotational orientations of the models. The 2-dimensional representations were projected down the fiber axes to represent 1-dimensional distributions perpendicular to the axis. The model projections had a pixel size of 0.1 nm.

(d) TMV image projection

The experimental TMV image intensities were averaged by projection of individual virus segments down the long axis. The orientation of each segment was refined by cross-correlation (Steven *et al.*, 1984) using the highly defocused image, which was recorded with high-dose and contrast. The orientation of the same segment in the minimally defocused image could not be refined directly, due to low contrast. Instead, the minimally defocused and highly defocused images of the same TMV segment were rotationally aligned in 2 steps. First, the films were pre-aligned by placing each film at the same location on the light box using a stationary plastic template. This pre-alignment was reproducible to better than 0.1°, but only accurate to about 2.0°, because of variations in the positions of the images on the film. Second, the 2 images were aligned by digital analysis of the edges of the images, which were about 0.5 cm inside the edge of the film, were straight and had high contrast. The edges of both images were digitized and cross-correlated to determine the relative rotation of the images. The reproducibility of the angular alignment was better than 0.1°. Thereafter, TMV segments from the same 2 images were digitized using the template, digitally rotated by the proper angle to achieve alignment and projected down the axis.

About 10% of the projections were discarded because of large deviations from mirror symmetry attributed to surface contamination. All remaining projections were aligned using cross-correlation and averaged. The combined projection represented an average of 55 independent TMV segments from a single image at 780 nm defocus (a combined length of ~4 μm). The resolution of the combined projection was estimated to be 1.9 nm using the differential phase residual method (Frank *et al.*, 1981). The resolution of the combined projection was estimated from the spatial frequency beyond which the phase error was >45°.

(e) Image compensation

The electron images represent certain spatial frequencies of the specimen better than others. In order to represent correctly all spatial frequencies, it is necessary to compensate for the modulation transfer functions of the microscope, film and digitization system. The microscope contrast transfer function (CTF) was estimated using the first-order theory of Erickson & Klug (1971).

The film transfer function was estimated from the analytical representation of Downing & Grano (1982). The recording transfer function of the CCD digitization was modeled by a sinc function appropriate to the pixel dimensions (Brigham, 1973). At f8.0, the photographic lens of the CCD system did not affect the image. Under our operating conditions, the film recording and digitization effects were small, predictable and easily compensated (Langmore & Smith, 1992). However, the CTF has a dominant role in imaging and varies dramatically between images. Compensation for the CTF is an unmet challenge, especially at low spatial frequencies.

In order to compensate for the CTF we needed to empirically determine the value of Q , the ratio of maximum scattering contrast to maximum phase contrast, which is the only free parameter in determining the microscope transfer function (Erickson & Klug, 1971). We used a simple method adapted from Toyoshima & Unwin (1988) to determine Q from 2 images of the same molecules at different defocus values. We first determined the ratio of the 1-dimensional equatorial Fourier transforms of the combined projections of TMV at 780 and 4620 nm defocus as a function of spatial frequency. The transforms were normalized at zero spatial frequency. We then used different values of Q (from 0.06 to 0.22) to calculate the ratio of the calculated CTF at both defocus values. Ratios from the observed amplitudes and CTF were compared at a spatial frequency of 0.1 nm^{-1} , where there was a peak in the empirical transform of TMV and a large ratio predicted between the CTFs at the 2 defocus values. A Q value of 0.14 gave the best agreement between the ratio observed in the TMV data and that predicted from the theoretical CTF. This value of Q was used for all image calculations and compensations shown in this paper.

In order to examine the theoretical effects of defocus, projections of TMV models were calculated at different defocus. Cylindrically averaged models were projected, Fourier transformed, multiplied by the theoretical CTF and reverse-transformed. Model resolution was reduced to that of the data (1.9 nm) by imposing a bell-shaped low-pass filter that had a value of unity for spatial frequencies less than 0.56 nm^{-1} , and a value of 0.01 at 0.77 nm^{-1} .

In order to compensate the observed image intensities, Fourier transforms of the experimental projections were divided by the theoretical CTF and reverse-transformed. The resulting compensated projection was converted from image intensity, $I(x)$, to scattering probability, $P(x)$, using the equation $P(x) = -\ln [I(x)/I_b(x)]$, where $I(x)$ is the projected image intensity (see above) and $I_b(x)$ is the estimated intensity of the ice background, interpolated from the image intensities surrounding the TMV. The

empirical scattering probabilities of different images can be compared, independent of ice thickness and incident electron intensity. All scattering projections and radial distributions shown represent absolute scattering probabilities, except as specifically noted.

(f) *Calculation of absolute mass, mass per unit length and mass density*

The absolute scattering probabilities can be used to estimate the mass per unit area, using the theoretical coefficient relating scattering cross-section to mass, i.e. the mass scattering coefficient, S (nm^2/amu). By integration over an area or a length, this information was used to estimate the total mass and mass per unit length of TMV. By radial reconstruction using the real-space algorithm described by Steven *et al.* (1984) the scattering probabilities were used to estimate the radial mass densities of TMV. If these estimates agree with the published values for TMV, we can be confident that we understand the imaging of molecules in ice, and that cryo-EM has the potential to accurately determine internal densities of biological molecules of unknown structure.

Table 1 is a summary of the theoretical mass scattering coefficients of ice and biological material in vacuum and in vitreous ice, containing averages over all the atoms of each type of molecule. The molecular cross-sections were calculated as described above and by Langmore & Smith (1992). The mass scattering coefficients were calculated by dividing the molecular cross-sections by the molecular masses. In all cases 70 to 80% of the scattering is inelastic. Of course the magnitude of the scattering coefficients in ice are substantially less than in vacuum, due to the exclusion of solvent from the molecular volumes. The mass scattering coefficients are almost independent of specimen composition in vacuum, but are specimen-dependent in ice owing to the differences in molecular density. In ice, elastic scattering from nucleic acids is predicted to be more than twice as great as from protein, whereas inelastic scattering is similar from the 2 molecules. Thus, energy filtration should not only greatly increase the probability of scattering from frozen-hydrated molecules, but also decrease the dependence upon molecular composition.

Table 1 also shows the predicted scattering amplitudes of protein and nucleic acid in vacuum and in ice, which show that the phase contrast (which is proportional to scattering amplitude) is the least dependent of specimen composition. From the predicted ratios on scattering cross-section to amplitude, it is expected that the ratio of scattering to phase contrast depends upon the physical nature of the specimen (Erickson & Klug, 1971).

Table 1
Theoretical mass scattering coefficients

| Molecule | Vacuum | | | | | Ice | | | |
|----------|---------------------------------------|-------------------------------------------------------------|---------------------------------------------------------------|--------------------------------------------------------------|--------------------------------------------------|-------------------------------------------------------------|---------------------------------------------------------------|--------------------------------------------------------------|--------------------------------------------------|
| | Density (g/cm^3) | $S_{\text{el}} \times 10^6$ (nm^2/amu) | $S_{\text{inel}} \times 10^6$ (nm^2/amu) | $S_{\text{tot}} \times 10^6$ (nm^2/amu) | $f(0) \times 10^2$ (nm/amu) | $S_{\text{el}} \times 10^6$ (nm^2/amu) | $S_{\text{inel}} \times 10^6$ (nm^2/amu) | $S_{\text{tot}} \times 10^6$ (nm^2/amu) | $f(0) \times 10^2$ (nm/amu) |
| Ice | 0.92 | 4.24 | 10.03 | 14.27 | 1.91 | — | — | — | — |
| Protein | 1.33 | 4.13 | 11.10 | 15.23 | 2.32 | 1.19 | 4.16 | 5.35 | 0.99 |
| NaRNA | 1.89 | 4.58 | 10.75 | 15.33 | 2.12 | 2.51 | 5.87 | 8.38 | 1.19 |
| NaDNA | 1.79 | 4.42 | 10.39 | 14.81 | 2.08 | 2.24 | 5.24 | 7.48 | 1.10 |

S_{el} , excess elastic mass scattering coefficient.

S_{inel} , excess inelastic mass scattering coefficient.

S_{tot} , total excess mass scattering coefficient.

$f(0)$, excess electron scattering amplitude at zero angle.

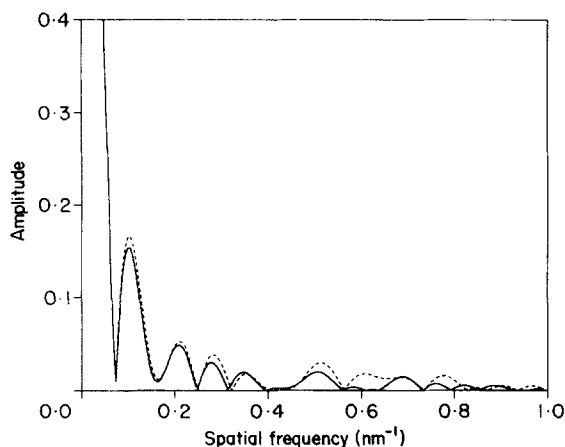


Figure 1. Comparison of the equatorial Fourier transform of the model at 1.0 nm resolution (----) with the equatorial X-ray scattering amplitudes calculated from Caspar (1956) (—). Comparison of the observed and predicted X-ray Fourier amplitudes gave a crystallographic R factor of 0.05 at 2.0 nm resolution and 0.1 at 1.0 nm resolution, where $R = \sum |F_{\text{obs}} - F_{\text{pred}}| / \sum F_{\text{obs}}$. Fourier amplitudes were normalized to unity at zero spatial frequency.

There are 3 potential sources of experimental error in our mass measurements: measurement of magnification; fiber tilt; and electron-induced mass loss. Use of the helical repeat of TMV to determine local magnification minimizes magnification error to <3%. Foreshortening of tilted TMV molecules would have increased the scattering per unit length. This effect would have been noticed, because tilt would make an apparent change in the magnification. Mass loss due to radiation damage could affect the electron scattering from TMV. At room temperature the mass loss from dry TMV is about 30% after a dose of 600 electrons/nm² at 70 keV (Lamvik & Langmore, 1977). At liquid nitrogen temperature the mass loss from dry TMV is about 1% after the same dose (Wall & Hainfeld, 1986). Mass loss for molecules in vitreous ice has never been carefully measured, but our preliminary results show less than a 5% loss in scattering from frozen-hydrated TMV during a dose of 3600 electrons/nm² (not shown).

3. Results

(a) Validity of the TMV modeling

To test the validity of the computer model, an electron density map of TMV was built and used to predict equatorial X-ray scattering in liquid water. Figure 1 compares the equatorial Fourier transform of the model with the equatorial X-ray scattering amplitudes calculated from Caspar (1956). The amplitudes agreed very well to 2.0 nm resolution, with a crystallographic R factor of 0.05, confirming that the TMV model is valid to at least 2.0 nm resolution. The R factor increased to 0.10 when spatial frequencies to 1.0 nm⁻¹ were included.

(b) Effects of resolution and defocus upon calculated electron images and reconstructions

Cryo-EM images were calculated at different resolution and defocus values in order to simulate the effects of the CTF on the images. The effect of

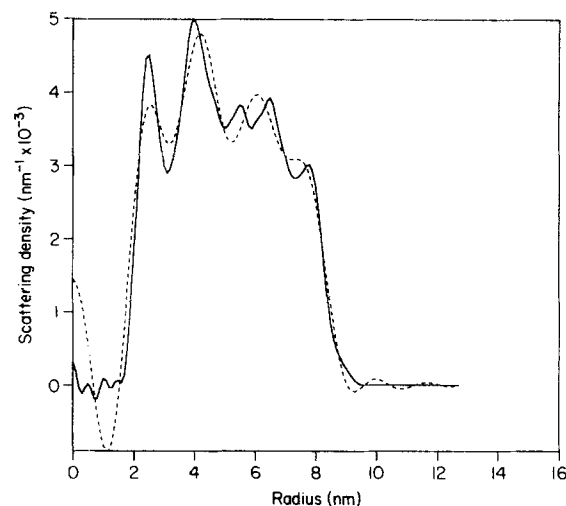


Figure 2. Predicted radial electron scattering density distributions of frozen-hydrated TMV at different resolution. Resolution limit is 0.5 nm (—) or 1.9 nm (----).

resolution is shown in Figure 2. At 0.5 nm resolution, the predicted TMV radial reconstruction has five distinct density maxima, and relatively flat solvent density at the center and edge. However, at 1.9 nm resolution, the reconstruction has only four peaks, due to the merging of two high-resolution maxima. As expected from the representation of the object by Bessel functions, the resolution limit most affects the center of the reconstruction, where a large spurious peak is created. Nevertheless, the low-resolution reconstruction has the major features of TMV. Further comparisons of the theoretical radial density distributions with the experimental reconstructions are at 1.9 nm resolution.

The effects of the CTF on the projections from model TMV images are demonstrated in Figure 3(a) to (d) for defocus values of -100, 30, 100 (Scherzer, 1949) and 780 nm. In each case, the first zero of the CTF is outside the resolution limit of 1.9 nm, yet minimal overfocus or underfocus substantially modifies the images. The effects are much more dramatic at the defocus values of 1 to 3 μm commonly used for cryo-EM.

The significant effects of the CTF are evident in the radial reconstructions of the TMV, shown in Figure 3(e) to (h). Even at the Scherzer (1949) defocus peak magnitudes are significantly altered. Therefore, to correctly obtain quantitative information from cryo-EM images, the images must be recorded very near focus, or the CTF must be compensated. Collection of data very close to focus is impractical in cryo-EM, because the very low image contrast (and hence low signal-to-noise ratio) make image analysis problematic. We conclude that the data must be recorded with significant defocus and compensated for the CTF.

(c) Accuracy of the compensated images and radial reconstructions

The ability to compensate the CTF by simple first-order theory is well known (Erickson & Klug,

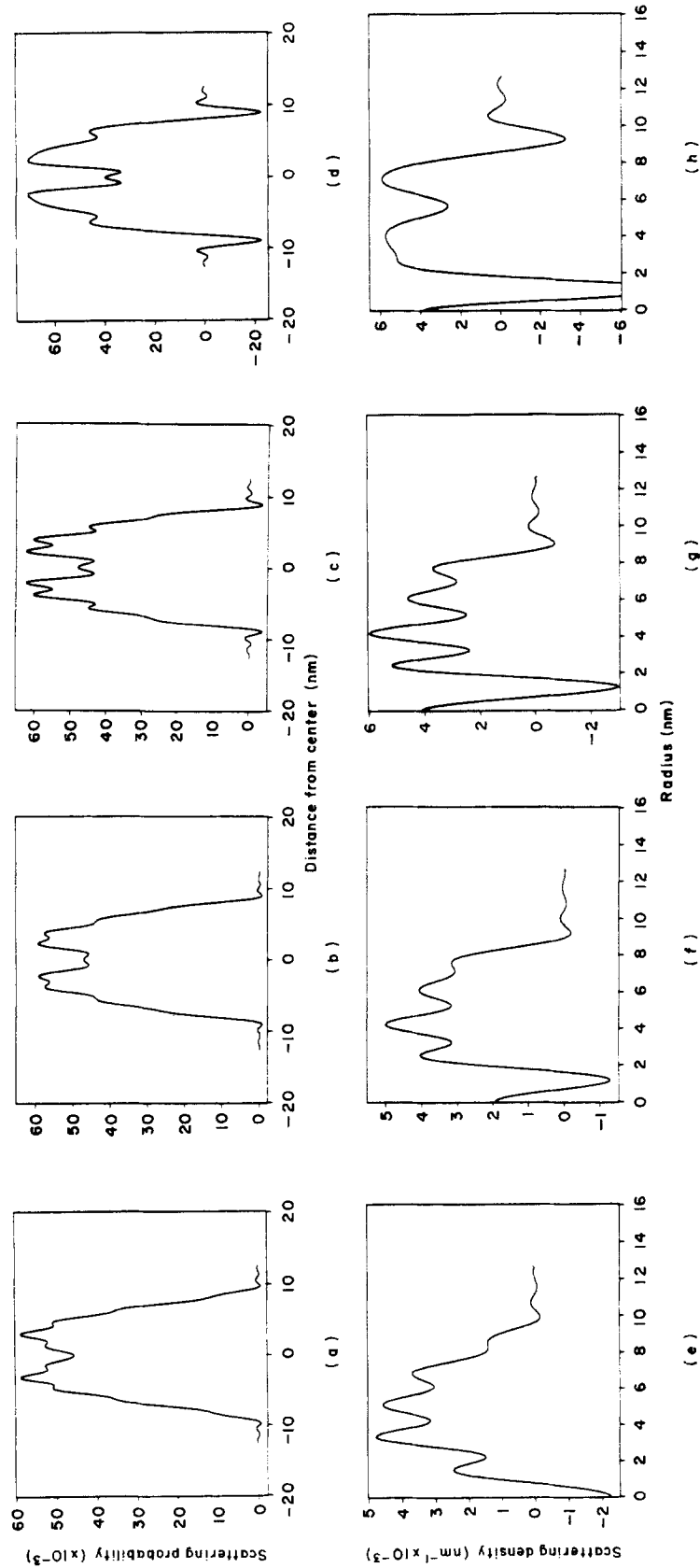


Figure 3. Predicted 1.9 nm resolution projections and reconstructions of frozen-hydrated TMV at different values of defocus. (a) to (d) Projected scattering probabilities at defocus values of (a) -100 nm; (b) 30 nm; (c) 100 nm (Scherzer); and (d) 780 nm. (e) to (h) Radial scattering densities calculated from the projections in (a) to (d).

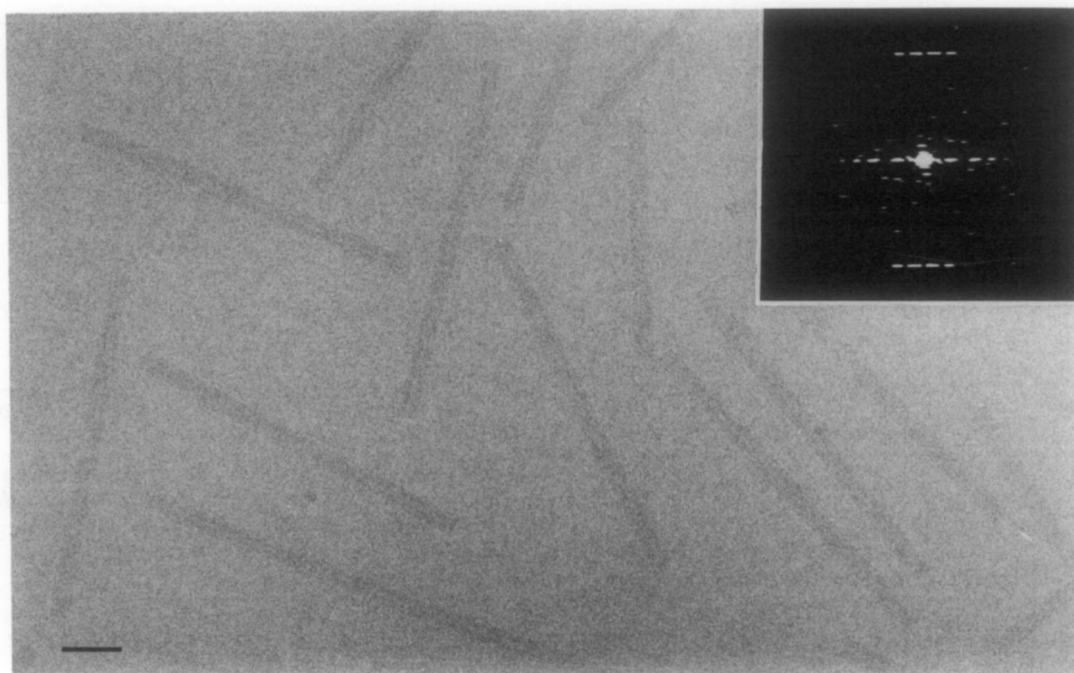


Figure 4. Low-dose bright-field image of energy-filtered frozen-hydrated TMV at 780 nm defocus. Inset is a calculated Fourier transform of a 100 nm segment of TMV. The bar represents 50 nm.

1971; Erickson, 1971). However, the fidelity of the compensated images has never been proven experimentally for biological molecules of known structure. Thus, it is important that the compensated TMV images be compared to the atomic structure of TMV. Typical TMV images at 780 nm defocus, and a representative Fourier transform are given in Figure 4. The strong helical transform confirms that the particles have remained well-ordered in both the equatorial and meridional directions.

The CTF at 780 nm defocus calculated from the first-order theory is given in Figure 5(a). The empirical value for the parameter Q was determined to be 0.14 by comparing the images at 780 and 4620 nm defocus (Materials and Methods). The equatorial Fourier amplitudes from the calculated and empirical cryo-EM images before and after CTF compensation were compared in order to test the validity of the first-order theory in predicting the CTF. Comparison of the Fourier transform of the model with that of the image at 780 nm defocus is given in Figure 5(b). The large differences between the relative peak heights illustrate the necessity for compensation. After compensation (Fig. 5(c)), the agreement between the model and experimental Fourier transform is excellent, with a crystallographic R factor of 0.12. The extremely low R factor reflects the accuracy of the model and scattering calculations, and the adequacy of the first-order theory to compensate the images from very low spatial frequencies to the resolution limit of 1.9 nm.

In order to test the sensitivity of image compensation to the details of compensation, the R factor was calculated after compensation with a range of values of Q , shown in Figure 6. The R factor is a minimum for $0.13 < Q < 0.17$, with substantial

degradation of the fit between model and compensated image outside that range.

Projected scattering probabilities of TMV before and after CTF compensation, and the predicted projection at the same resolution are given in Figure 7. There is little resemblance between the projection obtained from the defocused image data and that predicted theoretically. The projection obtained after compensation, however, closely resembles the theoretical projection, confirming that the compensated images correctly represent the molecular structure.

(d) *Absolute scattering probability and mass per unit length*

The data in Figure 7(b) can also be compared to determine how well the absolute scattering from TMV agrees with the predictions based on the model and the theoretical mass scattering coefficients that have been summarized in Table 1. The integrated scattering probabilities (scattering cross-section per unit length) were 0.75 and 0.73 nm for the cryo-EM data and model, respectively. Thus, the empirical mass scattering coefficient was only 2.5% greater than that predicted. Use of the predicted average mass scattering coefficients for protein and RNA given in Table 1 also results in an integrated scattering probability of 0.73 nm, assuming 95% protein, 5% RNA and a mass per unit length of 131.1 kDa/nm (2130 identical protein subunits of molecular weight 17,500; 3 ribonucleotides/subunit; 300 nm in length). Presumably, Table 1 could be used to determine the absolute mass, mass per unit length and mass densities of unknown biological molecules in vitreous ice.

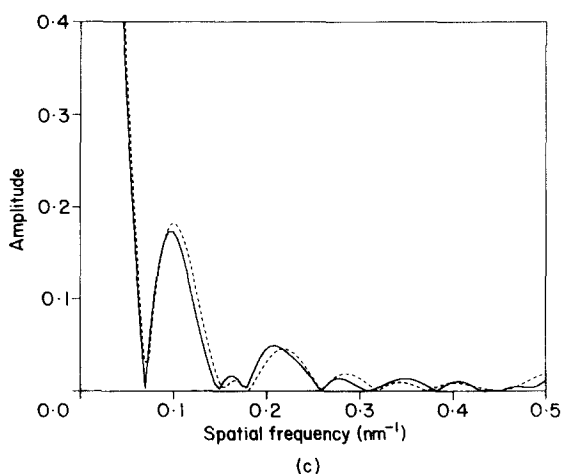
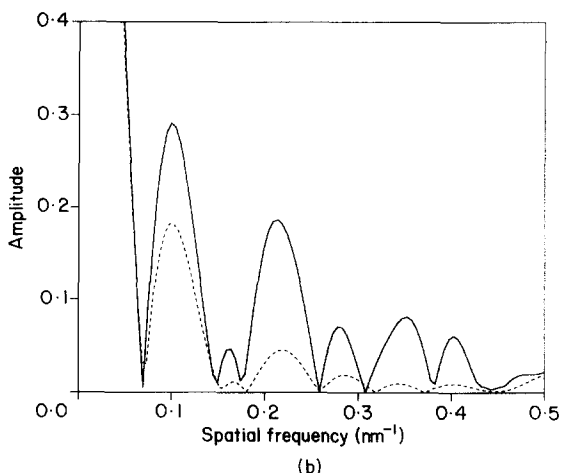
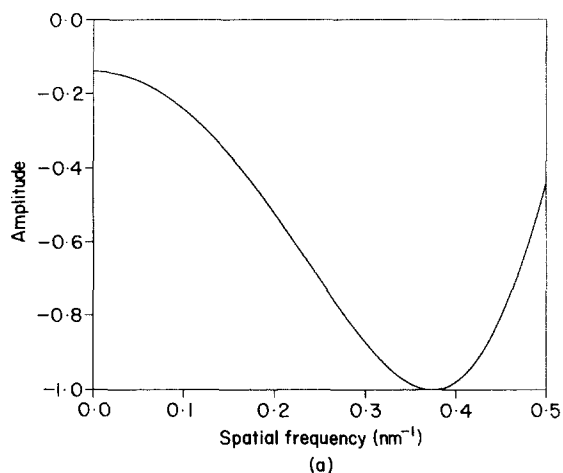


Figure 5. Observed and predicted equatorial Fourier amplitudes of frozen-hydrated TMV. (a) Calculated CTF at 780 nm defocus assuming $Q = 0.14$. (b) Fourier transform of the observed image with 780 nm defocus (—); Fourier transform of the predicted scattering from TMV (---). (c) Fourier transform of the observed image after CTF compensation (—); Fourier transform of the predicted scattering from TMV (---). The calculated R factor was 0.12. Fourier amplitudes were normalized to unity at zero spatial frequency.

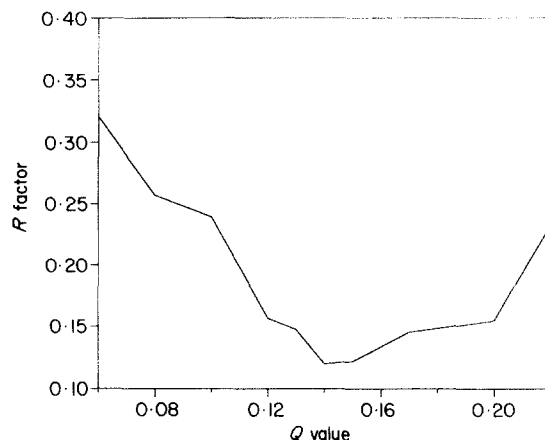


Figure 6. Sensitivity of image compensation to the choice of Q as measured by the crystallographic R factor. Fourier amplitudes were normalized to unity at zero spatial frequency and 1.9 nm resolution.

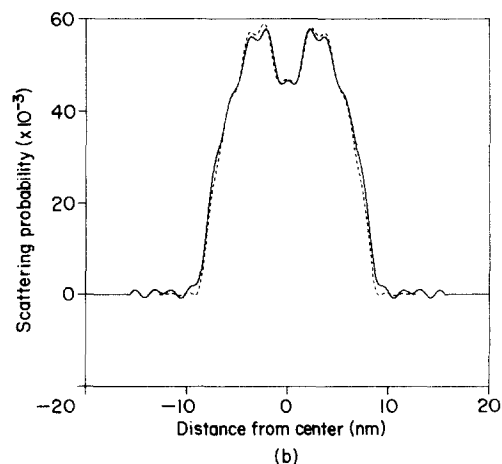
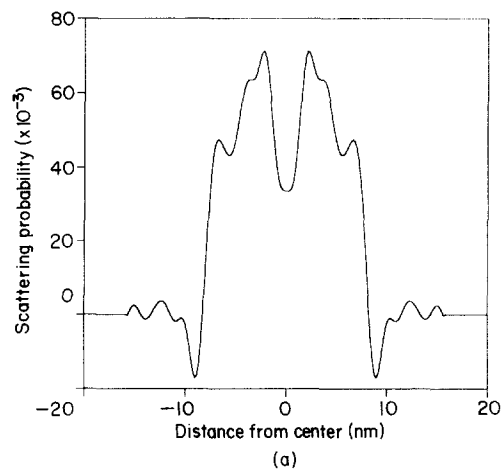


Figure 7. Effects of CTF compensation on projections of frozen-hydrated TMV. (a) Observed scattering probabilities of TMV at 780 nm defocus. (b) Comparison of the observed scattering probabilities after CTF compensation (—), and the predicted scattering probabilities (---) at 1.9 nm resolution.

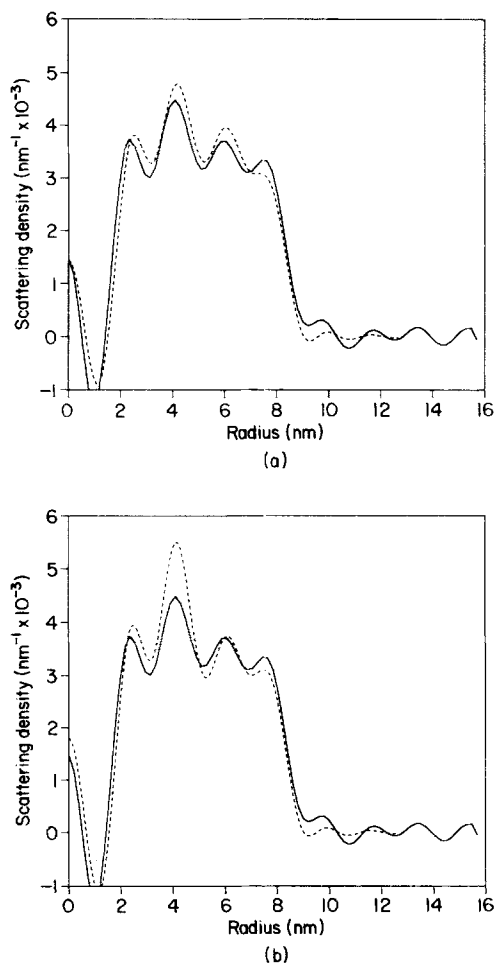


Figure 8. Differences between weighting atoms by scattering amplitude and cross-section. (a) Comparison of the observed radial scattering densities calculated from the CTF-compensated scattering probabilities in Fig. 7 (—), and the densities predicted by elastic scattering amplitude (---). (b) Comparison of the observed radial scattering densities calculated from the CTF-compensated scattering probabilities in Fig. 7 (—), and the distribution predicted by cross-section (---) at 1.9 nm resolution.

(e) *Cryo-EM density reconstructions*

The empirical scattering probabilities shown in Figure 7(b) were reconstructed into radial density distributions using the method of Steven *et al.* (1984), and are shown in Figure 8(a) and (b). The electron scattering density depends upon the sum of the phase contrast, elastic scattering contrast and inelastic scattering contrast, each of which has a slightly different dependence upon atomic composition (e.g. Table 1). Image compensation assumes that both the phase and scattering contrast from each atom type are proportional to the elastic scattering amplitude of that atom. We wondered whether the relative weighting of different atoms in the image are best predicted by elastic scattering amplitude (Fig. 8(a)) or total (elastic plus inelastic) scattering intensity (Fig. 8(b)). At 1.9 nm resolution, both theoretical and experimental distributions have peaks at 2.5, 4.2, 6.0, and 7.5 nm radius and have central holes of ~ 2 nm radius. The

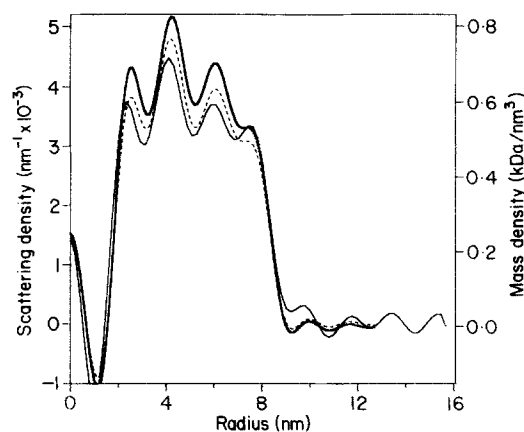


Figure 9. Comparison of predicted mass densities with experimental and predicted scattering densities of TMV at 1.9 nm resolution. Predicted radial mass densities (—); predicted radial scattering densities (---); CTF-compensated radial scattering densities (· · ·). The comparison of Fourier transforms of observed scattering densities and predicted mass densities gave an *R* factor of 0.12.

relative contributions of RNA and protein are best represented by the calculation based on scattering amplitudes, presumably due to the predominance of phase contrast. Thus, it seems better to use scattering amplitudes (as done throughout this paper) rather than the intensity of elastic and inelastic scattering to predict the images.

How are the reconstructions of the electron scattering densities related to the mass densities of TMV? Figure 9 compares the predicted radial mass densities, predicted radial scattering densities, and the experimental radial scattering densities. The peaks at 2.5, 4.2, 6.0 and 7.5 nm radius can be attributed to a series of reverse turns, the RNA, the LS, RS, RR, and LR helices, and the C helix (Namba *et al.*, 1989). The scattering densities from the compensated cryo-EM data appear to be a valid representation of the radial mass density distribution of both RNA and protein.

4. Discussion

(a) *Computer modeling of TMV*

The validity of our TMV model depends upon the accuracy of the atomic co-ordinates and the choice of solvent structure. Because the X-ray data were good to 0.29 nm (Namba *et al.*, 1989), the atomic co-ordinates should be capable of modeling our images at 1.9 nm resolution. However, because of extremely high temperature factors, the four C-terminal residues were not included in the published map (Namba *et al.*, 1989). The absence of these residues in the model should result in an underestimate of scattering density at the outer edge of the TMV where the C terminus is located (Namba *et al.*, 1989), and a 2% underestimate of the total scattering.

The assumption of uniform solvent density might also create errors in the model. Badger & Caspar (1991) found a 5% variation in solvent density at

the surface of insulin, based on the X-ray scattering at low resolution. It is possible that the extremely high R factors they found at low spatial frequencies were due to solvent density fluctuations, incorrect choice of the boundary between solvent and protein or other factors. The Langridge procedure for modeling solvent (Langridge *et al.* 1960) avoids the issue of the protein/solvent boundary, but assumes a uniform solvent density and knowledge about the molecular volumes. In principle, the protein and/or RNA excluded volumes in our model could be incorrect. The calculated partial specific volume of the TMV model was only 3% lower than the experimental value for TMV, showing that our choices for partial specific volumes were probably correct at room temperature. Our comparison of predicted and observed X-ray scattering from TMV was normalized at zero spatial frequency and revealed excellent agreement in the limited spatial frequency range from 0 to 0.5 nm^{-1} . This confirms the validity of our modeling the structure in liquid water at low resolution.

(b) Modeling energy-filtered cryo-EM images

Modeling of energy-filtered cryo-EM images relies on the ability to predict accurately the scattering amplitudes and intensities of the molecule and solvent. Prediction of the molecular scattering is susceptible to several errors. Although the large-angle elastic scattering should be accurately predicted, the small-angle scattering depends upon the long-range shielding of the nuclei by those electrons that respond to the chemical environment. Thus, the scattering amplitudes from solids at small angles might be considerably different from those from single atoms. In addition, the inelastic cross-sections are difficult to calculate, because they are sensitive to chemical effects (e.g. see Reimer & Ross-Messemer, 1989). We have used semi-empirical cross-sections for scattering from biological material and ice. Another potential difficulty in using the inelastic scattering contrast is the inherent delocalization of the inelastic events, which should begin to affect the data at a resolution between 1 and 10 nm (Kohl & Rose, 1985). The experiments of Isaacson *et al.* (1974) indicated an apparent resolution of 1 to 2 nm in the inelastic images from carbon, but there are no reliable data on these effects.

Modeling of the solvent is another potential source for error. The success in modeling X-ray scattering in liquid water does not necessarily validate the modeling of density, structure or excluded volume of vitreous ice. In the absence of data on the density of thin layers of vitreous ice, we used the density measured for thick layers of vapor-deposited vitreous ice (Seats & Rice, 1982). Preliminary results indicate that there are few structural differences between vapor-deposited ice and the thin films of hyperquenched ice used for cryo-EM (Hallbrucker *et al.*, 1991). It is also possible that vitreous ice might be significantly more dense around biological molecules. Finally, we have not taken into

account the temperature dependence of the excluded solvent volume. The unit cell dimensions of purple membranes are decreased by about 2% during cooling to 100 K (Zaccari, 1987). A 2% decrease in each dimension of the TMV would result in a 6% increase in protein density.

Regardless of the potential difficulties in modeling biological molecules in vitreous ice, as described above, our results demonstrate that the absolute scattering from TMV can be predicted to within 3% of the experimental values. Either all of our assumptions are accurate or there are compensating errors.

(c) CTF compensation of energy-filtered cryo-EM images

The majority of electron microscope applications are limited to low resolution, where the density information is useful to delineate molecular boundaries and discriminate between different components, such as solvent, lipid, protein and nucleic acid. We have shown that these low resolution applications require compensation for both the phase and scattering contrast contributions to the CTF.

Erickson & Klug (1971) were the first to propose a method to experimentally estimate the relative contributions of scattering and phase contrast, expressed as a value for the variable Q . They calculated two-dimensional Fourier transforms of thin catalase crystals and evaluated the amplitudes of specific reflections for a range of defocus values. By fitting the results to the expectations of the first-order theory, they estimated a Q value of 0.35 for a specimen stained with uranyl acetate. Toyoshima & Unwin (1988) used a similar technique and determined a Q value of 0.07 for frozen-hydrated acetylcholine receptors in 30 to 50 nm of vitreous ice.

We have adapted these techniques for use on single particles. The ratio of amplitudes in the continuous Fourier transforms of the combined projections at two defocus values were compared to the theoretical ratios predicted by the first-order theory. Using this method on TMV in 150 nm of ice, we calculated a value of 0.14 for Q . We have confirmed that compensation of the image with this value of Q gives an accurate reconstruction. The R factor between the compensated image and the model was a minimum for $Q = 0.14$. Q could be varied by ± 0.02 , without causing a significant error in the reconstruction, whereas larger changes in the value chosen for Q gave reconstructions that poorly represented internal densities.

The fact that our value of Q was twice that measured by Toyoshima & Unwin (1988) could be the result of our use of energy filtration or of thicker ice. Because of chromatic aberration, even the unfiltered bright-field EM is able to utilize a portion of inelastic scattering contrast. Without chromatic aberration, the inelastic scattering from the molecule would be perfectly focused back into the image of the molecule, giving no contrast. With

chromatic aberration, the inelastically scattered electrons are always defocused, moving a fraction of their number to different places in the image, creating scattering contrast. The fraction of the inelastic scattering that is utilized in the image will depend upon spatial frequency and defocus. The use of energy filtration insures that all of the inelastic scattering is used to form contrast, thus increasing the amount of scattering contrast relative to that found without energy filtration. In addition, Q should depend upon ice thickness for energy-filtered images (and for unfiltered images as well). The dependence upon ice thickness will result from the fact that in thicker ice the increased amount of inelastic scattering and multiple elastic scattering reduces phase contrast. This can be seen as failure of the "weak-phase-weak-amplitude" approximation.

(d) *Radial density analysis and mass quantitation by cryo-EM*

Quantitation of mass and mass density by cryo-EM has certain advantages over bulk solution methods such as low-angle X-ray scattering, which is a traditional way to measure the average mass and mass per unit length in aqueous environments. Bulk solution methods are subject to misinterpretation due to errors in measurement of concentration, specimen aggregation, molecular flexibility, and heterogeneity in size, shape and mass. In contrast, EM imaging allows individual molecules to be measured and relationships to be established between different parameters. For example, molecules having uniform diameters, high axial ratios, and no observable aggregation can be selected for study, and correlations determined between parameters such as size and molecular weight, diameter and mass per unit length, morphological features and density, etc.

Until recently, STEM has offered clear advantages over the conventional electron microscope for quantitative studies of unstained biological molecules. The high contrast and high signal-to-noise ratios of dark-field imaging made STEM ideal for quantitation of total mass and mass per unit length. However, dehydration of the specimens has limited STEM resolution to 2 to 4 nm, thereby making analysis of internal densities problematic. Unfortunately, frozen-hydrated molecules cannot be imaged with significant contrast in STEM, due to the difficulty in forming phase contrast. To obtain high-resolution images with significant contrast, it is necessary to use bright-field cryo-EM. Potential applications are the measurement of molecular weight, mass per unit length and three-dimensional densities of proteins, nucleic acids and lipids. Relatively high contrast and high signal-to-noise ratios can be obtained from low-dose defocused cryo-EM images, with further improvements possible by energy filtration.

STEM has been used to determine the radial density distribution of frozen-dried TMV, by selection of molecules that appeared well preserved by

the criterion of agreement between the expected and observed internal densities (Steven *et al.*, 1984), or by selection of all images with reasonable mirror symmetry (Smith *et al.*, 1990). Neither peak positions nor intensities of those reconstructions agree with our TMV model (Langmore & Smith, 1992), presumably due to distortions during freeze drying. Lepault (1985) used a phase-contrast bright-field image to determine the radial density distribution of frozen-hydrated TMV. That reconstruction did not resemble the predicted scattering density distribution presented here, presumably due to the fact that no corrections for the CTF were performed. In another study of frozen-hydrated TMV, Jeng *et al.* (1989) did not show or use the low spatial frequency information from the images due to their inability to compensate for the CTF. Their density maps, derived from X-ray scattering and electron images, seem to over-emphasize density at high radius compared to the X-ray density maps.

Our cryo-EM analyses of frozen-hydrated TMV have shown that compensation for the CTF results in a density reconstruction that agrees quantitatively with the atomic model and achieves higher resolution than the STEM analysis. It was also shown that the scattering densities from the CTF-compensated cryo-EM data accurately represent the mass density of protein and nucleic acid. Thus, we are confident that CTF compensation of experimental images can lead to accurate reconstructions of the absolute mass densities of frozen-hydrated proteins and nucleic acids. Although our experiments focused on one-dimensional reconstructions of the density as a function of radius, the conclusions are applicable to reconstructions in three dimensions.

We thank Dr G. Stubbs, Vanderbilt University, Nashville for the TMV, Dr D. Caspar for permission to reproduce the X-ray scattering data, Dr M. Ludwig, University of Michigan, Ann Arbor for use of the Silicon Graphics IRIS 4D/220GTX, and Mr G. Tank for determination of mass loss of TMV. We also thank Drs D. Caspar, R. Henderson, and C. Correll for useful discussions. These studies were supported by NIH GM27937, NSF DIR8706052 and NSF DMB9106659.

References

- Badger, J. & Caspar, D. (1991). Water structure in cubic insulin crystals. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 622-626.
- Brakenhoff, G. (1974). Relative mass determination in darkfield microscopy. In *Principles and Techniques of Electron Microscopy Biological Applications* (Hayat, M., ed.), vol. 4, pp. 16-44, VanNostrand Reinhold Co., New York.
- Brigham, E. O. (1973). In *The Fast Fourier Transform* (Oppenheim, A. V., ed.), Prentice Hall, NJ.
- Cammarano, P., Romeo, M., Gentile, A., Felsani, A. & Gualerzi, C. (1972). Size heterogeneity of large ribosomal subunits and conservation of the small subunits in eucaryote evolution. *Biochim. Biophys. Acta*, **281**, 597-624.
- Caspar, D. (1956). Radial density distribution in the

- tobacco mosaic virus particle. *Nature (London)*, **177**, 928-930.
- Chiu, W. (1986). Electron microscopy of frozen, hydrated biological specimens. *Annu. Rev. Biophys. Chem.* **15**, 237-257.
- Downing, K. & Grano, A. (1982). Analysis of photographic emulsions for electron microscopy of two-dimensional crystalline specimens. *Ultramicroscopy*, **7**, 381-404.
- Dubochet, J., Lepault, J., Freeman, R., Berriman, J. A. & Homo, J.-C. (1982). Electron microscopy of frozen water and aqueous solutions. *J. Microsc.* **128**, 219-237.
- Dubochet, J., Adrian, M., Chang, J.-J., Homo, J.-C., Lepault, J., McDowell, A. W. & Schultz, P. (1988). Cryo-electron microscopy of vitrified specimens. *Quart. Rev. Biophys.* **21**, 129-228.
- Erickson, H. (1971). The Fourier transform of an electron micrograph—first order and second order theory of image formation. In *Advances in Optical and Electron Microscopy*, vol. 5, pp. 163-199. Academic Press, New York.
- Erickson, H. & Klug, A. (1971). Measurement and compensation of defocusing and aberrations by Fourier processing of electron micrographs. *Phil. Trans. Roy. Soc. ser. B*, **261**, 105-118.
- Frank, J., Verschoor, A. & Boublik, M. (1981). Computer averaging of electron micrographs of 40S ribosomal subunits. *Science*, **214**, 1353-1355.
- Hallbrucker, A., Mayer, E., O'Mard, L. P., Dore, J. C. & Chieux, P. (1991). Structural characterization of hyperquenched glassy water and vapour-deposited amorphous ice. *Phys. Letters*, **A159**, 406-410.
- Isaacson, M., Langmore, J. & Rose, H. (1974). Determination of the non-localization of the inelastic scattering of electrons by electron microscopy. *Optik*, **41**, 92-96.
- Jeng, T.-W., Crowther, R., Stubbs, G. & Chiu, W. (1989). Visualization of α -helices in tobacco mosaic virus by cryo-electron microscopy. *J. Mol. Biol.* **205**, 251-257.
- Kohl, H. & Rose, H. (1985). Theory of image formation by inelastically scattered electrons in the electron microscope. *Advan. Electron. Electron Phys.* **65**, 175-200.
- Lamvik, M. & Langmore, J. (1977). Determination of particle mass using scanning transmission electron microscopy. In *Proc. 10th Annu. Scanning Electron Microsc. Symp.* p. 401, IIT Res. Inst. Chicago.
- Langmore, J. & Athey, B. (1987). Removal of inelastically scattered electrons substantially increases phase contrast on frozen-hydrated molecules. In *Proc. 45th Annual EMSA Meeting* (Bailey, G. W. ed.), pp. 252-253. San Francisco Press, San Francisco.
- Langmore, J. & Smith, M. (1992). Quantitative energy-filtered microscopy of biological molecules in ice. *Ultramicroscopy*, in the press.
- Langmore, J., Smith, M. & Rankert, D. (1990). Quantitative energy-filtered cryo-electron microscopy of hydrated chromosome fibers, viruses, and heavy atoms. In *Proc. XIIIth International Congress for Electron Microscopy* (Peachey, L & Williams, D., eds), pp. 652-653. San Francisco Press, San Francisco.
- Langridge, R., Marvin, D., Seeds, W. & Wilson, H. (1960). The molecular configuration of deoxyribonucleic acid. II. Molecular models and their Fourier transforms. *J. Mol. Biol.* **2**, 38-64.
- Lepault, J. (1985). Cryo-electron microscopy of helical particles TMV and T4 polyheads. *J. Microsc.* **140**, 73-80.
- Lepault, J. & Leonard, K. (1985). Three-dimensional structure of unstained frozen-hydrated extended tails of bacteriophage T4. *J. Mol. Biol.* **182**, 431-441.
- Namba, K., Pattanayek, R. & Stubbs, G. (1989). Visualization of protein-nucleic acid interactions in a virus: refined structure of intact tobacco mosaic virus at 2.9 Å resolution by X-ray fiber diffraction. *J. Mol. Biol.* **208**, 307-325.
- Reimer, L. & Ross-Messemer, M. (1989). Contrast in the electron spectroscopic imaging mode of a TEM. I. Influence of zero-loss filtering on scattering contrast. *J. Microsc.* **155**, 169-182.
- Sceats, M. & Rice, S. (1982). In *Water* (Franks, F., ed.), vol. 7, chapt 2, Plenum, New York.
- Schafer, L., Yates, A. & Bonham, R. (1971). New values for the partial wave scattering factor for the elements $1 \leq Z \leq 90$ for incident electron energies of 10, 40, 70, and 100 keV. *J. Chem. Phys.* **55**, 3055.
- Scherzer, O. (1949). The theoretical resolution limit of the electron microscope. *J. Appl. Phys.* **20**, 20-29.
- Schröder, R., Hofmann, W. & Ménétret, J.-P. (1990). Zero-loss energy filtering as improved imaging mode in cryoelectronmicroscopy of frozen-hydrated specimens. *J. Struct. Biol.* **105**, 28-34.
- Smith, M., Athey, B., Williams, S. & Langmore, J. (1990). Radial density distribution of chromatin: evidence that chromatin fibers have solid centers. *J. Cell Biol.* **110**, 245-254.
- Steven, A., Hainfeld, J., Trus, B., Steinert, P. & Wall, J. (1984). Radial distributions of density within macromolecular complexes determined from dark-field electron micrographs. *Proc. Nat. Acad. Sci., U.S.A.* **81**, 6363-6367.
- Stewart, M. & Vigers, G. (1986). Electron microscopy of frozen-hydrated biological material. *Nature (London)*, **319**, 631-636.
- Taylor, K. & Glaeser, R. (1974). Electron diffraction of frozen, hydrated protein crystals. *Science*, **106**, 1036-1037.
- Taylor, K. & Glaeser, R. (1976). Electron microscopy of frozen hydrated biological specimens. *J. Ultrastruct. Res.* **55**, 448-456.
- Toyoshima, C. & Unwin, N. (1988). Contrast transfer for frozen-hydrated specimens: determination from pairs of defocused images. *Ultramicroscopy*, **25**, 279-292.
- Toyoshima, C. & Unwin, N. (1990). Three-dimensional structure of the acetylcholine receptor by cryoelectron microscopy and helical image reconstruction. *J. Cell Biol.* **111**, 2623-2635.
- Wall, J. & Hainfeld, J. (1986). Mass mapping with the scanning transmission electron microscope. *Annu. Rev. Biophys. Biophys. Chem.* **15**, 355-376.
- Wall, J., Isaacson, M. & Langmore, J. (1974). The collection of scattered electrons in dark field electron microscopy. II. Inelastic scattering. *Optik*, **39**, 359-374.
- Zaccai, G. (1987). Structure and hydration of purple membranes in different conditions. *J. Mol. Biol.* **194**, 569-572.
- Zamyatin, A. (1972). Protein volume in solution. *Prog. Biophys. Mol. Biol.* **24**, 107-123.
- Zeitler, E. & Bahr, G. (1957). Contribution to the quantitative interpretation of electron microscope pictures. *Exp. Cell Res.* **12**, 44-65.
- Zeitler, E. & Bahr, G. (1962). A photometric procedure for weight determination of submicroscopic particles by quantitative electron microscopy. *J. Appl. Phys.* **33**, 847-853.