ultramicroscopy

Magnification mismatches between micrographs: corrective procedures and implications for structural analysis

A. Aldroubi [a,1], B.L. Trus [b,c], M. Unser [a], F.P. Booy [c] and A.C. Steven [c]

[a] *Biomedical Engineering and Instrumentation Program*, [b] *Computer Systems Laboratory, Division of Computer Research and Technology*, [c] *Laboratory of Structural Biology Research, National Institute of Arthritis, Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892, USA*

Quantitative structural analysis from electron micrographs of biological macromolecules inevitably requires the synthesis of data from many parts of the same micrograph and, ultimately, from multiple micrographs. Higher resolutions require the inclusion of progressively more data, and for the particles analyzed to be consistent to within ever more stringent limits. Disparities in magnification between micrographs or even within the field of one micrograph, arising from lens hysteresis or distortions, limit the resolution of such analyses. A quantitative assessment of this effect shows that its severity depends on the size of the particle under study: for particles that are 100 nm in diameter, for example, a 2% discrepancy in magnification restricts the resolution to ~5 nm. In this study, we derive and describe the properties of a family of algorithms designed for cross-calibrating the magnifications of particles from different micrographs, or from widely differing parts of the same micrograph. This approach is based on the assumption that all of the particles are of identical size: thus, it is applicable primarily to cryo-electron micrographs in which native dimensions are precisely preserved. As applied to icosahedral virus capsids, this procedure is accurate to within 0.1-0.2%, provided that at least five randomly oriented particles are included in the calculation. The algorithm is stable in the presence of noise levels typical of those encountered in practice, and is readily adaptable to non-isometric particles. It may also be used to discriminate subpopulations of subtly different sizes.

1. Introduction

The technique of three-dimensional reconstruction from sets of projection images of (intrinsically alike) particles is becoming an increasingly effective and widely practiced method of structural analysis [1,2]. Such procedures have particularly high potential in applications to cryo-electron micrographs of particles suspended in thin films of vitreous ice, for which the preservation of native structure is optimal [3-5]. In consequence, the mutual compatibility of the data is high, and the prospects of extending the analysis to relatively high resolution are enhanced. Current reconstructions of icosahedral virus capsids typi-cally incorporate data from 20-40 particles, and achieve resolutions of 3.0-4.0 nm (e.g., refs. [6-9]). If possible, all particles in a given reconstruction are extracted from a single micrograph, and accordingly, were imaged under almost identical conditions. However, extension to substantially higher resolution will require combining larger numbers (e.g., several hundreds) of particles, which will make the combining of data from several different micrographs inevitable.

Although disparities of magnification are by no means the only complication that arises when data from different micrographs are to be combined, they are a significant factor, particularly in studies that aspire to relatively high resolution. Owing primarily to lens hysteresis [10,11], the magnifications of micrographs recorded at nominally the same setting may vary by a few percent. Moreover, lens distortions [10] may result in mag-

---

[1] To whom correspondence should be addressed: Rm. 3W13, Bldg. 13, BEIP/NIH, Bethesda, MD 20892, USA.

nification variations of the order of one percent over the field covered in a single micrograph. We have made a quantitative assessment of the resolution-limiting effects of magnification mismatches on structural analyses (appendix A). These effects may be considered as damping the specimen's Fourier transform, to an extent that becomes progressively more severe at higher spatial frequencies, and is systematically worse for larger particles. However, if magnifications can be standardized to within 0.2%, resolutions of up to 0.4 nm are accessible for particles up to 100 nm in diameter, before this effect becomes a significant problem.

Thus, it is desirable that the sampling rates of all digital images to be combined in a given analysis be standardized to within a few tenths of a percent. In this paper, we present an empirical approach that is capable of determining the relative scaling of "spherical" particles, such as icosahedral virus particles, to within this margin of error. Standardized sampling may then be imposed by interpolation. Algorithms capable of performing this calibration are derived, several of their major properties are proved, and the specifics of their implementation are summarized. Their performance is then illustrated in model experiments involving computer-generated data, both in the absence and in the presence of noise. Next, we present a number of applications to cryo-electron micrographs of icosahedral virus capsids. These include examples of particles viewed in differing orientations, particles imaged at differing values of defocus and particles from widely separated parts of the same micrograph. Finally, we discuss adaptations of this procedure to handle non-spherical particles, as well as powder diffraction patterns calculated from micrographs.

2. Mathematical theory and algorithms

Before addressing the problem of finding the relative magnifications of particles in one or several micrographs, we first derive an algorithm for one-dimensional curves. To solve the matching problem for the particles in micrographs, we will

apply this algorithm to their radial autocorrelation functions. In this section, we state the mathematical problem and derive a solution. The proofs are deferred to appendix B.

2.1. Statement of the problem

Problem A: Find the scalar stretch factor s between the two functions f and g, given that:

$$g(t) = f(st), \qquad \forall t \in [0, T]. \tag{1}$$

Problem B: Find the scalar stretch factor s and the amplification factor M between the two functions f and g, given that:

$$g(t) = Mf(st), \qquad \forall t \in [0, T]. \tag{2}$$

In practice, f and g are not aligned, and they are corrupted by noise. Thus, we state two related problems that are more realistic.

Problem A': Find the scalar stretch factor s between the two functions f and g, given that:

$$g(t) = f(st - c) + \eta(t), \qquad \forall t \in [0, T], \tag{3}$$

where $\eta(t)$ denotes a random noise component and c is an unknown constant.

Problem B': Find the scalar stretch factor s and the amplification factor M between the two functions f and g, given that:

$$g(t) = Mf(st - c) + \eta(t), \qquad \forall t \in [0, T]. \tag{4}$$

Our aim is to derive accurate, robust, and fast algorithms to solve problems A' and B'.

2.2. The averaging kernels

Without loss of generality, we consider only positive functions. We start with problems A and B, and assume that $T = \infty$. Since our aim is to find algorithms that are not sensitive to noise, we will consider averaging schemes. If we average eq. (1) using a kernel $K(t)$, and perform a change of variable, we obtain:

$$\int_0^\infty K(t) \, g(t) \, dt = \int_0^\infty K(t) \, f(st) \, dt$$

$$= \frac{1}{s} \int_0^\infty K(t/s) \, f(t) \, dt. \tag{5}$$

If the kernel is separable in the sense that

$$K(t/s) = H(t) L(1/s), \qquad \forall t,s > 0, \qquad (6)$$

then we can use eq. (5) to find the stretch factor s as follows:

$$s = N^{-1} \left[\int_0^\infty K(t) \, g(t) \, dt \Big/ \int_0^\infty H(t) \, f(t) \, dt \right], \qquad (7)$$

where $N^{-1}(s)$ is the functional inverse of $N(s)$ which is defined to be:

$$N(s) = L(1/s)/s. \qquad (8)$$

This motivates us to find separable kernels, i.e., kernels satisfying eq. (6). We can characterize all such functions, a fact that we state in the following theorem:

Theorem 1. If $K(x) \in C^1(0, \infty)$, $K(x) > 0$, and satisfies the separability condition:

$$K(xy) = H(x) L(y), \qquad \forall x,y > 0,$$

for some functions $H(x)$ and $L(x)$, then

$$K(x) = ax^\alpha,$$

where $a > 0$ and α are arbitrary constants.

We note that with this type of kernel, eq. (5) takes the form of a Mellin transform [24] evaluated at a fixed value α.

2.3. Algorithm for problem A

For problem A, we have that $g(t) = f(st)$. As before, we assume that all of our functions are positive, and without loss of generality, we assume that the scaling factor s is such that $0 < s \le 1$. If we choose a kernel $K(t) = at^\alpha$, then solution (7) takes the simple form:

$$s = \left(\int_0^\infty t^\alpha f(t) \, dt \Big/ \int_0^\infty t^\alpha g(t) \, dt \right)^{1/(\alpha+1)}. \qquad (9)$$

In practice, however, the interval $[0, T]$ on which our functions are defined, is finite. In this case, eq. (7) becomes an implicit nonlinear equation to be solved for s. Taking $T = 1$, we obtain:

$$s = \left(\int_0^s t^\alpha f(t) \, dt \Big/ \int_0^1 t^\alpha g(t) \, dt \right)^{1/(\alpha+1)}. \qquad (10)$$

Eq. (10) is highly nonlinear because the integral in the numerator is nonlinear in s, and also because of the power $(\alpha + 1)^{-1}$. An analytic solution cannot be derived, but we can use an iterative scheme to solve for s. For this purpose, we define the function $R_\alpha(x)$ by:

$$R_\alpha(x) = \left(\int_0^x t^\alpha f(t) \, dt \Big/ \int_0^1 t^\alpha g(t) \, dt \right)^{1/(\alpha+1)}. \qquad (11)$$

To find s, we start with an initial estimate (x_0) of the value of s and compute a new value $x_1 = R_\alpha(x_0)$. By iteration, we converge to the value of s, a result stated in the following theorem:

Theorem 2. The equation,

$$x_{n+1} = R_\alpha(x_n), \qquad (12)$$

has a steady state $\bar{x} = s$. Moreover, if the condition:

$$\int_0^s t^{\alpha+1} f'(t) \, dt < 0, \qquad (13)$$

where $f'(t)$ is the derivative of $f(t)$, then $\bar{x} = s$ is asymptotically stable.

Condition (13) is a decay condition stating that the function f must (on average) be decaying. If f is, in fact, decreasing, then $f'(t)$ is negative, and condition (13) is always satisfied.

2.4. Algorithm for problem B

For problem B, we have that $g(t) = Mf(st)$ in the interval $[0,1]$. Again, we assume that all of our functions are positive and, without loss of generality, that the scaling factor s is such that $0 < s \le 1$. We proceed as above to obtain a system of two nonlinear algebraic equations having (s, M) as solution:

$$s = M^{1/(\alpha-1)} \left(\int_0^s t^\alpha f(t) \, dt \Big/ \int_0^1 t^\alpha g(t) \, dt \right)^{1/(\alpha+1)};$$

$$M = s^{\beta+1} \left(\int_0^1 t^\beta g(t) \, dt \Big/ \int_0^s t^\beta f(t) \, dt \right), \qquad (14)$$

where $\beta < \alpha$ are chosen to be positive, but are otherwise arbitrary. To solve eq. (14), we first define the function $Q_\beta(y)$ by:

$$Q_\beta(y) = \left(\int_0^1 t^\beta \, g(t) \, dt \Big/ \int_0^y t^\beta \, f(t) \, dt \right). \qquad (15)$$

Using eqs. (11) and (15), we now define a convergent iterative scheme which yields the solution (s, M). This process is described in the next theorem.

Theorem 3. The system of difference equations:

$$\begin{cases} x_{n+1} = y_n^{1/(\alpha+1)} R_\alpha(x_n), \\ y_{n+1} = x_n^{\beta+1} Q_\beta(x_n), \end{cases} \qquad (16)$$

has a steady state $(\bar{x}, \bar{y}) = (s, M)$.
Moreover, if β is chosen so that $\beta < \alpha$ and the conditions:

$$\int_0^s t^{\alpha+1} f'(t) \, dt < 0, \qquad \int_0^s t^{\beta+1} f'(t) \, dt < 0, \qquad (17)$$

hold, then the resulting solution (s, M) is asymptotically stable.

It is important to note that, because of continuity, the two conditions in eq. (17) can be reduced to the first condition alone, provided that we choose β sufficiently close to α.

2.5. Algorithm for problems A' and B'

The decay conditions (13) and (17) are not necessarily satisfied by real signals. Moreover, in practice, the data are noisy and the functions f and g need not be in exact registration. Accordingly, we cannot apply the algorithms (12) or (16) directly. Instead, we apply the algorithms to the autocorrelation functions $A_f(t)$ and $A_g(t)$ of f and g, respectively. The autocorrelation function of a signal $h(t)$ is by definition:

$$A_h(t) = \int_{-\infty}^{+\infty} h(\tau) \, h(t+\tau) \, d\tau. \qquad (18)$$

Usually, autocorrelation functions of real data are essentially decaying, and satisfy the convergence conditions of the algorithms. This is be-

cause the autocorrelation function of a signal has a maximum at the origin, and because data points tend to be more correlated to their closest neighbors than to their distant ones. If $g(t) = Mf(st)$, then the relation between $A_f(t)$ and $A_g(t)$ is given by:

$$A_g(t) = \frac{M^2}{s} A_f(st). \qquad (19)$$

For the case of zero-mean white noise that is independent of the data, as in problems A' and B', the relation between $A_f(t)$ and $A_g(t)$ is:

$$A_g(t) = \frac{M^2}{s} A_f(st) + E^2 \delta(t), \qquad (20)$$

where $\delta(t)$ is the Dirac distribution and E^2, the noise energy.

Since the algorithms use kernels of the form $k(t) = t^\mu$ with μ a positive number, the Dirac delta function in eq. (20) has no effect on the result because it is multiplied by $k(0) = 0$. Thus, the algorithms are insensitive to independent white noise. Moreover, we note that eq. (20) does not depend on the constant c in eqs. (3) and (4).

2.6. Implementation for micrographs

The implementation of the procedure that determines the scaling factor s and the contrast factor M relating two micrographs (or two distinct locations within a single micrograph) F and G proceeds as follows:

(i) Choose sets of particles from F and G, and use them to calculate azimuthally averaged 1D autocorrelation functions A_f and A_g. Store A_f and A_g in vectors of dimension D.

(ii) Choose values for α and β with $\alpha < \beta$ (e.g., $\alpha = 1$ and $\beta = 0.1$). Compute the four integrals that appear in $R_\alpha(x)$ and $Q_\beta(x)$ in eqs. (11) and (15) on the discrete grid points $x = mi$ with $i = 1, \cdots, D$ where $m = (D)^{-1}$ (D = dimension of the vectors A_f and A_g). This is done by a quadrature formula such as the trapezoidal approximation or Simpson's rule for integration. Store the results in 1D arrays $W_{1\alpha}$, $W_{2\alpha}$, $W_{1\beta}$, $W_{2\beta}$.

(iii) Start with an initial guess (e.g., $(x_0, y_0) = (1.0, 1.0)$) and find a value i_0 such that $Mi_0 \leq x_0 \leq M(i_0 + 1)$.

(iv) For each 1D array $W_{1\alpha}$, $W_{2\alpha}$, $W_{1\beta}$, $W_{2\beta}$, use the components i_0 and $i_0 + 1$ to obtain, by linear interpolation, the values of the integrals appearing in $R_\alpha(x_0)$ and $Q_\beta(x_0)$.

(v) Compute x_1 and y_1 from x_0 and y_0 using the difference eqs. (16).

(vi) Repeat the last three steps (i.e., (iii)–(v)) using x_1 and y_1 as the new estimate.

3. Results

3.1. Model experiments with computer-generated data

(i) *Convergence properties of the algorithm.* To evaluate our algorithms, several tests were first performed with computer-generated data. In the experiment illustrated in fig. 1, the two one-dimensional curves (the continuous line and the dashed line) represent the same function, but differ in sampling rate by the factor ($s = 0.98$), and in amplitude by the factor ($M = 1.2$). This test function is smoothly varying and monotonically decreasing, and thus satisfies condition (17). In these respects, it resembles the one-dimensional autocorrelation functions that we use in practice with spherical particles (see below).
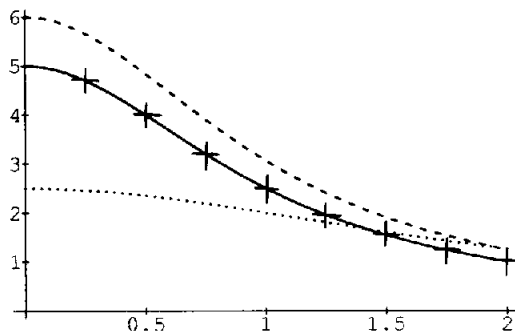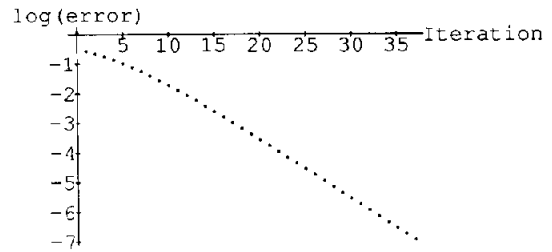
Fig. 2. Error in the scaling factor s versus the number of iterations of the algorithm.

Starting from the rather bad initial guess of ($s = 0.5$; $M = 5.0$) – see fig. 1 – the algorithm converges below the pre-set threshold of 0.1% tolerable error after 37 iterations. The rate of convergence, plotted in fig. 2, is such that the residual error decays in a quasi-exponential manner.

Fig. 3 displays a function that is not strictly decreasing, although it does show an overall tendency to decline over the range covered, and for ($\alpha = 1$; $\beta = 0.1$), condition (17) is satisfied. The correct scaling factors ($s = 0.97$; $M = 1.2$) were found to within 0.1% in 21 iterations, starting from ($s_0 = 0.9$; $M_0 = 0.99$), cf. fig. 3.

These experiments indicate that, as expected on theoretical grounds (section 2), the algorithm converges stably, and quite rapidly, despite the fact that poor initial guesses were used. In practice, the initial estimates for both s_0, assessed
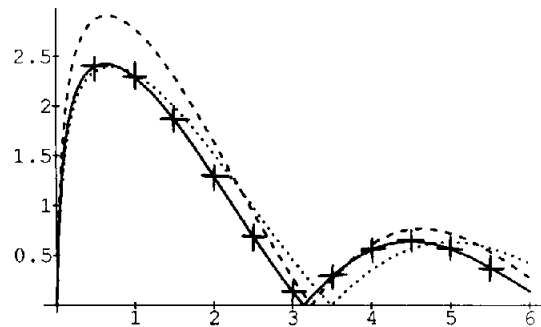
Fig. 1. Test function (continuous line) $f(x) = 5(1 + x^2)^{-1}$. Modified function $g(x) = Mf(sx)$, $s = 0.98$, $M = 1.2$ (dashed line). Initial guess $s_0 = 0.5$, $M_0 = 0.5$ (dotted line). Function found after 37 iterations ($\alpha = 1$, $\beta = 0.1$) for which the errors in s and M are below the pre-set threshold of 0.1% (crosses).

Fig. 3. Test function (continuous line) $f(x) = |3(0.1 + x)^{-1} \sin(x)|$. Modified function $g(x) = Mf(sx)$, $s = 0.97$, $M = 1.2$ (dashed line). Initial guess $s_0 = 0.9$, $M_0 = 0.99$ (dotted line). Function found after 21 iterations ($\alpha = 1$, $\beta = 0.1$) for which the errors in s and M are below the pre-set threshold of 0.1% (crosses).

from measurements of particle diameter made by manually applying a cursor to monitor-displayed particles, and M_0, effected by imposing some standard prescription for image normalization (e.g., refs. [12,13]) are unlikely to be in error by more than a few percent. Accordingly, in practice, convergence is likely to occur faster than these model experiments suggest.

(ii) *Robustness in the presence of noise.* To test how noise might affect the stability of the algorithm, an experiment was carried out using a triangular signal contaminated with noise. The results showed that the algorithm was completely stable up to signal-to-noise ratios of the order of one. In practice, it is applied to heavily averaged signals, in which the noise initially present in the micrograph has been largely suppressed, first by forming the two-dimensional autocorrelation function, then by azimuthal averaging and finally by averaging over a number of particles. Accordingly this experiment indicates that the performance of the algorithm should be oblivious to the low levels of residual noise (see also (iii) below).

3.2. The algorithms in practice: applications to cryo-electron micrographs

(iii) *Micrographs of herpes simplex virus capsids.* We have performed a number of experiments with such electron micrographs [9]. From single-particle images, (e.g., fig. 4a), the autocor-

Table 1

The relative magnifications of two sets of HSV capsid images from the same cryo-electron micrograph were adjusted to bear the ratio 0.9794 by interpolation; the algorithm was then used to estimate this factor in the presence of increasing amounts of added random Gaussian noise (i.e., noise in addition to that already present in the micrograph); this experiment was performed with $\alpha = 1$ and $\beta = 0.3$

SNR (added noise)	S calculated	Error (%)
infinity	0.9791	0.03
7:5	0.9801	0.07
1:1	0.9802	0.08
2:3	0.9751	0.44
1:2	0.9728	0.67

relation function was calculated (fig. 4b), and then azimuthally averaged (fig. 4c). The uncorrelated noise gives a spike at the origin which makes a negligible contribution to the integrals in eq. (14) since it is suppressed by factor t^α. Initially, 9 HSV capsids, representing a random distribution of orientations, were selected for analysis. Their ACFs were averaged, and then compared to the average ACF of the same nine images after resampling them with a relative scaling factor of 0.9794. Table 1 summarizes the results given by the algorithm both without any added noise, and with Gaussian white noise added to the rescaled images. With no added noise, the algorithm estimated the calibration factor to
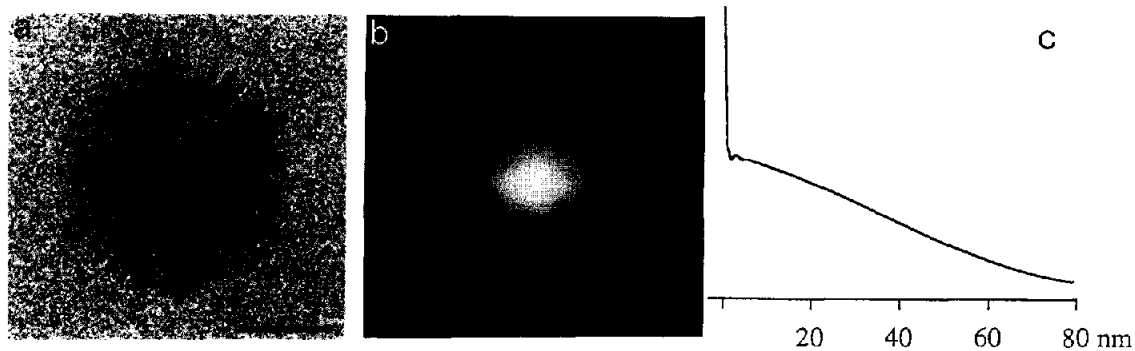

Fig. 4. (a) Electron micrograph of a herpes simplex virus nucleo capsid. (b) The autocorrelation function of the image. (c) The radial autocorrelation function.

within 0.03% error. The error increased to 0.67% when the total dynamic power of the added noise was twice that of the original images (SNR = 1 : 2).

(iv) *Significance of parameters α and β.* Eqs. (11) and (15) define a whole family of algorithms, in which α and β enter as free parameters. Accordingly, they may be assigned values that are likely to optimize the performance of the algorithm. We examined how the solution of the same experiment described in (iii) above (but without added noise) responded to changes in the values of these parameters. The results (table 2) suggest that smaller values of α and β give more accurate results, and that a ratio of $\alpha : \beta \gg 1$ is to be preferred. To choose higher values of α would assign disproportionate weight to the autocorrelation function values at outer radii, whereas when low values of α are used, comparable significance is assigned to the whole radial range.

(v) *Icosahedral particles in different orientations.* Icosahedral virus capsids vary considerably in the extent to which they depart from sphericity. For example, rotavirus [8] and reovirus [14] are almost perfectly spherical, whereas adenovirus [15] and iridovirus [16] have relatively flat, planar, facets. In this respect, herpesvirus capsids are intermediate between these extremes [9,17, 18]. Although differences in the projection that arise from viewing geometry tend to be smoothed out when the two-dimensional autocorrelation function is calculated (fig. 4b), one might expect

Table 3

This table summarizes the results of experiments performed to examine how the calibration factor given by the algorithm varies from particle to particle in consequence of differing viewing geometries of the HSV capsid; it also shows how such fluctuations are suppressed by including increasing numbers of randomly oriented particles in the calculation

Number of particles	Average error (%)	Min. error (%)	Max. error (%)	Number of trials
1	0.63	0.11	1.6	10
2	0.31	0.03	0.76	5
3	0.21	0.03	0.51	5
4	0.11	0.04	0.18	5
5	0.09	0.01	0.16	5

the scaling result to be somewhat sensitive to particle orientation, and that such dependence would be more pronounced for more angular polyhedral capsids. To probe this effect, we compared a number of herpesvirus capsid projections in different orientations, which had previously been determined. These images were taken from the same region of the same micrograph, and we assume that the particles are identical in size. The algorithm's output indicates that anomalous variations in scaling can result from differing particle orientations (table 3). With herpesvirus capsids, these fluctuations are typically at the 0.6% level, but exceptionally, may range as high as 1.6%. We also tested how the algorithm's performance might be stabilized by averaging over a number of randomly chosen particles. The results (table 3) indicate that stable results of the desired precision may be expected when at least five randomly oriented particles are averaged in each case.

(vi) *Micrographs recorded at different defocus values.* Phase contrast effects in cryo-micrographs [4,5] generate interference ("Fresnel") fringes at the particles' edges, the depth, width, and number of which vary as a function of defocus. In principle, therefore, one might expect significant differences in defocus between micrographs to register as apparent differences in magnification, i.e., anomalous size differences between identical particles. To investigate this effect, we compared a set of ten particles, as represented in three micrographs recorded sequentially in a focal se-

Table 2

The same experiment as that described in the legend to table 1 was performed, except that the values assigned to parameters α and β were varied; the signal-to-noise ratio was chosen to be one; the results are rather stable for values such that $1 < \alpha < 1.4$ and $0.3 < \beta < 0.8$

α	β	S calculated	error (%)
1	0.3	0.9802	0.08
1	0.4	0.9802	0.08
1	0.5	0.9803	0.09
1	0.6	0.9804	0.10
1	0.8	0.9805	0.11
1.2	0.3	0.9803	0.09
1.4	0.3	0.9804	0.10
2	1	0.9809	0.15
3	1	0.9811	0.17

rics, in which the first zero of the contrast transfer function occurred at $(1.6 \text{ nm})^{-1}$, $(2.4 \text{ nm})^{-1}$, and $(4.2 \text{ nm})^{-1}$, respectively. Since the same particles are compared, their sizes and orientations are identical in each micrograph. However, small but significant differences in apparent size are detected by the algorithm (table 4). We conclude that, when matching micrograph magnifications in practice, care should be taken to work either with images that were recorded at quite similar defocus settings, or with micrographs which have been adequately "restored" to compensate for phase contrast effects. Of course, data taken from micrographs with substantially different defocus values have more basic inconsistencies than a slight apparent mismatch of magnification.

(vii) *Variations in magnification across a given field.* We have used our algorithms to investigate the extent to which variations in magnification may occur across the field of view covered in a single micrograph. Such variations arise primarily from lens distortions [10] and, at a practical level, depend on how close the C2 lens is to crossover. In this experiment, the average one-dimensional autocorrelation function was calculated from clusters of five particles each, taken from the center and from each corner of a single micrograph recorded at a magnification of 36000. The results are shown in fig. 5. They reveal systematic variations in magnification in excess of 1% between opposite corners of the micrograph. This conclusion was confirmed with a similar analysis
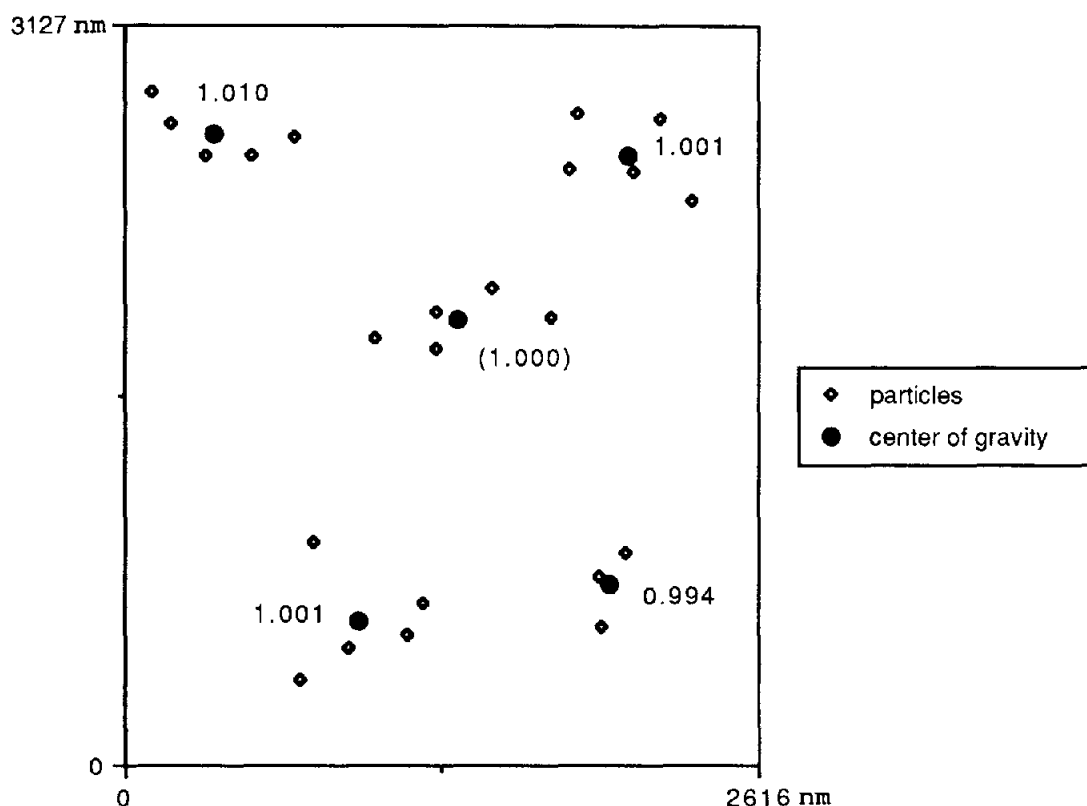

Fig. 5. This diagram represents the outcome of an experiment in which the scaling algorithm was used to probe for magnification variations between different parts of the same image. This cryo-electron micrograph was recorded at a magnification of 36000×, so that the area covered is 2.3 by 2.8 μm. Clusters of HSV capsids in the four corners (each site is marked with a cross) were compared with a cluster in the center. With five particles per cluster, the uncertainty in each measurement should be < 0.1% (table 2).

of a second micrograph, which revealed disparities in magnification that were somewhat smaller but of the same order (data not shown). Such excursions are tolerable for analyses of relatively small particles, but become significant for analyses of particles in the 100 nm range, in which resolutions of ≤ 3 nm are hoped for (see appendix A).

Thus, detectable size variations can occur within a given micrograph as a consequence of lens distortions (real variations) or focal gradients (anomalous variations). The presence of focal gradients may be detected by comparing the optical diffractograms from each corner of the micrograph.

(viii) *Improved three-dimensional reconstructions from combined data from different micrographs mutually scaled by this algorithm*. In one herpesvirus capsid reconstruction [B.L. Trus, F.P. Booy, W.W. Newcomb, J.C. Brown and A.C. Steven – unpublished results], the particles present in the micrograph analyzed were found not to be randomly oriented, and so these data did not give a uniform sampling of the icosahedral asymmetric unit (cf. [6]). The resulting reconstruction (not shown) was very noisy compared with a reconstruction performed with a full cover-
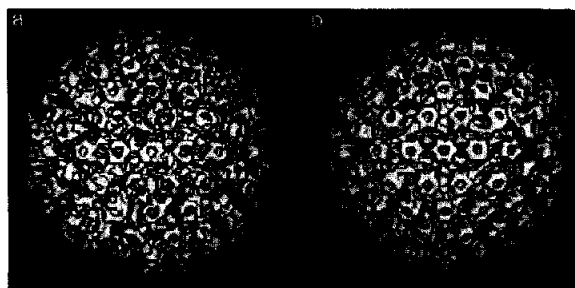

Fig. 6. (a) Reconstruction where the scaling factors are found manually. (b) Reconstruction where the scaling factors are found by the algorithm.

ing of the three-dimensional Fourier transform. We attempted to improve this reconstruction by including other, appropriately oriented, particles from another micrograph with a similar defocus value. At this stage in the analysis, the two micrograph magnifications were scaled on the basis of manual measurements of the diameters of particles on each. Little improvement was registered in the quality of the reconstruction (fig. 6a). However, when the procedure was repeated using the magnification scaling factor determined by this algorithm, a considerable improvement was obtained (fig. 6b). The algorithm's result showed that the initial manual estimate of relative magnification had been in error by $\sim 1.5\%$. This seemingly small discrepancy was enough to have a substantial effect on the resulting reconstruction.

4. Discussion

4.1. Principle of the method

Our method is based on the assumption that the particles analyzed are identical in size, and that it is therefore valid to use them to cross-calibrate magnifications, both between micrographs and within a given micrograph. This assumption is well founded in the case of cryo-electron micrographs of icosahedral virus capsids, whose dimensions are precisely determined in the assembly process, and are preserved upon embedding in vitreous ice [19]. We believe that this also holds true for cryo-electron micrographs of

Table 4

Differences in defocus lead to anomalous differences in apparent magnification; this table summarizes the calibration factors given for the same set of ten particles imaged in a focal series in which the three exposures (I, II, III) corresponding to the first zero of the contrast transfer function being at $(1.6\ \mathrm{nm})^{-1}$, $(2.4\ \mathrm{nm})^{-1}$, and $(4.2\ \mathrm{nm})^{-1}$, respectively

Micrographs		S calculated	error (%)
Reference	To be scaled		
II	I	0.9932	0.68
III	I	1.0018	0.18
III	II	1.0087	0.87

Note that the *apparent* size does not change monotonically with defocus value. However, the scaling factors measured in this experiment combine consistently, viz. $s_{I \to III} = s_{I \to II} \times s_{II \to III}$. The particles' structure is likely to change somewhat in response to increased electron dose over the focal series. However, the fact that the consistency condition is met indicates that this effect on apparent magnification is slight compared to that of defocus-induced changes in phase contrast, at least over the limited dose range covered.

other kinds of macromolecules and supramolecular complexes. With non-spherical particles, however, projected dimensions are orientation-dependent, so that use of the algorithm is more involved than straightforward analysis of the autocorrelation function (see below).

With specimens prepared for electron microscopy by conventional methods, distortions caused by dehydration, flattening, staining, etc., may change the particles' projected dimensions by several percent or more, so that the condition necessary for applicability of our method (uniform size) is not met. For such data, we do not see a way to improve on the traditional methods of magnification calibration, such as optical diffraction analysis of micrographs of catalase, or some other periodic test specimen. We note, however, that it has been reported that magnification discrepancies of the order of 2–5% among negatively stained specimens may be detected by correspondence analysis [25].

4.2. Size dependence

Quantitative appraisal of the implications of combining data with different magnifications of a three-dimensional reconstruction shows that this effect manifests itself as a progressive damping of the structure factors in the composite three-dimensional Fourier transform towards higher spatial frequencies (appendix A). Specification of the resolution limit for particles of a given size, as imposed by a particular level of disparity in magnification, may be made with reference to fig. 7 and table 5. The effect is systematically more pronounced for larger particles. This trend repre-
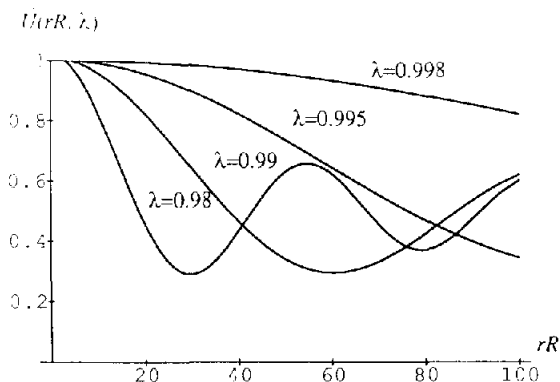

Fig. 7. The radial dependence of the magnification mismatch-associated damping function (eq. (A.5)), as averaged over all azimuths, is plotted for several different relative scaling factors (λ). The independent variable is the product (rR), where r is the radial coordinate in real space, and R is the radial coordinate in Fourier space.

sents a second obstacle to achieving a given resolution for larger particles, as compared with smaller particles, since it has already been demonstrated that larger particles require a correspondingly higher number of independent projections [20].

4.3. Applicability

In applications to cryo-electron micrographs of icosahedral capsids, in order to attain the desired accuracy of within 0.1–0.2% in magnification calibration, the following conditions must be satisfied: (1) At least five particles in random orientations, and preferably quite close together, must be taken from each micrograph. (2) The micro-

Table 5
The resolution limits (R_0). expressed in reciprocal space units, imposed by magnification mismatches for particles of different radii (r_0) are compiled for several different degrees of mismatch (λ); for instance, $\lambda = 0.98$ is equivalent to a 2% mismatch; the radius attributed to purple membrane (crystalline bacteriorhodopsin) is the radius required to cover its hexagonal unit cell

| | R_0 | | | |
	$\lambda = 0.98$ (2%)	$\lambda = 0.99$ (1%)	$\lambda = 0.995$ (0.5%)	$\lambda = 0.998$ (0.2%)
HSV capsid $r_0 = 62.5$ nm	$(4.71 \text{ nm})^{-1}$	$(2.35 \text{ nm})^{-1}$	$(1.18 \text{ nm})^{-1}$	$(0.47 \text{ nm})^{-1}$
Bacteriophage T7 capsid $r_0 = 25$ nm	$(1.88 \text{ nm})^{-1}$	$(0.94 \text{ nm})^{-1}$	$(0.47 \text{ nm})^{-1}$	$(0.188 \text{ nm})^{-1}$
Purple membrane $r_0 = 3.5$ nm	$(0.264 \text{ nm})^{-1}$	$(0.132 \text{ nm})^{-1}$	$(0.066 \text{ nm})^{-1}$	$(0.026 \text{ nm})^{-1}$

graphs should have approximately the same degree of defocus, or have been correctly restored to standardize phase contrast effects.

In practice, three-dimensional density maps may be refined by using an initial reconstruction to generate the two-dimensional projection corresponding to each constituent view, and then applying the scaling algorithm to calibrate their precise magnifications. After appropriate adjustment of the scaling of each particle, a refined reconstruction is then calculated.

4.4. Detectability of size differences

Under certain circumstances, a preparation may contain particles of several different (discrete) sizes. In the case of polymorphism of a given virus capsid, the size differences should be sufficiently great that the respective polymorphs may be readily distinguished, and analyzed separately. More subtle differences may arise in the event of co-isolation of two or more viruses that are similar in size, and morphologically alike. However, it seems reasonable to expect that the presence of multiple components with size differences as small as 1% or so may be detected by systematic use of our algorithm.

4.5. Non-spherical particles

With cryo-electron micrographs of non-spherical particles, variants of this procedure may be employed. In these cases, the particles present in each image should first be analyzed by correlation alignment, classification, and averaging [13,21,22]. Then, line scans across corresponding averaged images from the different micrographs under comparison should be processed with our algorithm, instead of the azimuthally averaged autocorrelation functions used here. The results obtained with noise in simulated experiments (section 3) suggest that the algorithm is sufficiently robust to perform well under these conditions, particularly if preceded by a significant amount of averaging. An alternative approach to applying the algorithm to calibrate the magnifications of particles is to apply it to the azimuthally averaged powder diffraction pattern calculated for a set of such particles. With regular particles that contain repetitive internal structure, such structure may show up as rings in the powder pattern (see, for example, fig. 5 of ref. [9]), which provide a suitable signal to which the algorithm may be applied.

Acknowledgements

Appendix A. Resolution-limiting effects of magnification mismatches when micrographs are combined

When the information contents of two or more electron micrographs are to be combined, this process may be performed in two dimensions by averaging them after correlation alignment, or in three dimensions by some reconstruction algorithm, after determining their relative orientations. In either case, the act of combining them may be viewed as an averaging operation. If (assuming perfect mutual alignment) we consider corresponding points on the respective particles, one is at radius r, and the other at λr, i.e., they are mutually displaced by a distance $d = r(1 - \lambda)$. Thus, one may expect details that are smaller than d to be blurred out in the averaged image. This effect is negligible close to the origin (i.e., near the center of the particle), but grows linearly with increasing radius. Thus, it is evident that the problem becomes more acute for large particles (which have larger outer radii) than for small particles.

More generally, we may consider this effect as a progressive attenuation of the particles' composite Fourier transform towards higher spatial frequencies, i.e., at relatively high resolution. In

two dimensions, the two particles' Fourier transforms in polar coordinates are given by:

$$F_1(R, \phi) = \int_0^{+\infty} \int_0^{2\pi} f(r, \theta)\, e^{-j2\pi rR \cos(\phi - \theta)}$$
$$\times 2\pi r\, dr\, d\theta, \qquad (A.1)$$

$$F_2(R, \phi) = \int_0^{-\infty} \int_0^{2\pi} f(\lambda r, \theta)\, e^{-j2\pi rR \cos(\phi - \theta)}$$
$$\times 2\pi r\, dr\, d\theta, \qquad (A.2)$$

where R is the radial frequency and ϕ the Fourier angle. Changing variable r in eq. (A.2), and averaging them, we obtain:

$$\tfrac{1}{2}(F_1(R, \phi) + F_2(R, \phi))$$
$$= \int_0^{+\infty} \int_0^{2\pi} f(r, \theta)\, e^{-j2\pi rR \cos(\phi - \theta)}$$
$$\times U(r, R, \theta, \phi, \lambda) 2\pi r\, dr\, d\theta, \qquad (A.3)$$

where

$$U(r, R, \theta, \phi, \lambda)$$
$$= \tfrac{1}{2}\left(1 + \frac{1}{\lambda^2} e^{-j2\pi rR \cos(\phi - \theta)(\frac{1}{\lambda} - 1)}\right). \qquad (A.4)$$

This gives essentially the same quantity as eq. (A.1), apart from the additional factor U. Note that $0 \leq |U| \leq 1$; i.e., U is a damping factor.

To obtain a more balanced estimate of the overall effect of this distortion, we compute a radial damping factor by taking the average over all angles, which yields

$$\overline{U}(r, R, \lambda) = \frac{1}{2\pi} \int_0^{2\pi} U(r, R, \theta, \phi, \lambda)\, d\theta$$
$$= \tfrac{1}{2}\left\{1 + \frac{1}{\lambda^2} J_0[2\pi rR(1/\lambda - 1)]\right\}, \qquad (A.5)$$

where $J_0(x)$ is Bessel function of order 0, which is defined by

$$J_0(x) = \frac{1}{2\pi} \int_0^{2\pi} \exp[-jx \cos(\phi - \theta)]\, d\theta. \qquad (A.6)$$

Note that U is a function of the product $p = rR$. This function $\overline{U}(r, R, \lambda)$ is plotted in fig. 7 for several pertinent values of λ. The resolution limit

is determined by searching for the radial frequency R_0 at which the damping becomes significant, assuming that the particle is confined within a radius r_0. The cut-off limit is arbitrarily set to $1/\sqrt{2}$. Hence, we find that the radial cut-off frequency and particle radius are inversely proportional to each other:

$$R_0 r_0 = p(\lambda), \qquad (A.7)$$

where the critical value $p(\lambda)$ is the solution of the equation $\overline{U}(p, \lambda) = 2^{-1/2}$. This technique was used to obtain estimates for the resolution limits for three representative cases of particles (cf. table 5).

We conclude that for relatively small particles (i.e., $r_0 \approx 10$ nm), no significant effect arises until resolutions that are high by current standards (viz. 0.1 nm^{-1}). However, with large particles, such as HSV capsids, the effects of magnification mismatches at the 1–2% level are already substantial, even at the moderate resolutions of 0.3–0.4 nm^{-1} that are currently accessible.

Finally, we note that the formalism presented above is equally applicable to quantitative assessment of other resolution-limiting effects, e.g., translational and orientational displacements, arising from imperfect alignment. In each case, the corresponding displacement operation is substituted in eq. (A.2), and the analysis followed through in otherwise the same way.

Appendix B. Proofs of theorems

B.1. Proof of theorem 1

Since

$$K(xy) = H(x)\, L(y), \qquad \forall x, y > 0, \qquad (B.1)$$

we have that:

$$H(x)\, L(y) = H(y)\, L(x), \qquad \forall x, y > 0. \qquad (B.2)$$

We use eq. (B.2) and group the terms to get:

$$\frac{L(y)}{H(y)} = \frac{L(x)}{H(x)}, \qquad \forall x, y > 0. \qquad (B.3)$$

Since x and y are independent we have that:

$$\frac{L(y)}{H(y)} = \frac{L(1)}{H(1)}, \qquad \forall y > 0. \tag{B.4}$$

Using eq. (B.1) together with eq. (B.4) we conclude that:

$$K(xy) = \lambda K(x) K(y), \qquad \forall x, y > 0, \tag{B.5}$$

where λ is some positive constant.

We differentiate eq. (B.5) with respect to x and evaluate at $x = 1$:

$$yK'(1y) = \lambda K'(1)K(y). \tag{B.6}$$

Eq. (B.6) is a differential equation that we solve to obtain:

$$K(y) = cy^{\alpha}, \tag{B.7}$$

where $\alpha = \lambda K'(1)$ and c is an arbitrary positive constant.

B.2. Proof of theorem 2

The steady-state solutions of the difference equation (12) must satisfy

$$\bar{x} = R_{\alpha}(\bar{x}). \tag{B.8}$$

The value $\bar{x} = s$ satisfies eq. (B.8) since by assumption we have that $g(t) = f(st)$. In order to study the stability of the steady state, we linearize at the steady state. To do this, we need to evaluate the derivative of $R_{\alpha}(x)$ at the steady state $\bar{x} = s$:

$$R'_{\alpha}(s) = \frac{s^{\alpha+1}}{\alpha+1} \frac{f(s)}{\int_0^s t^{\alpha} f(t) \, dt}. \tag{B.9}$$

A simple integration by part yields:

$$R'_{\alpha}(s)$$
$$= \frac{s^{\alpha+1} f(s)}{\alpha+1}$$
$$\times \left(\frac{s^{\alpha-1} f(s)}{\alpha+1} - \frac{1}{\alpha+1} \int_0^s t^{\alpha+1} f'(t) \, dt \right)^{-1} \tag{B.10}$$

If eq. (13) is satisfied then it follows from eq. (B.10) that $0 < R'_{\alpha}(s) < 1$ which implies the asymptotic stability of the steady state.

B.3. Proof of theorem 3

Because of eq. (14), (s, M) is a steady state of system (16). In order to study its stability, we linearize the right side at the steady state of eq. (16) and obtain the Jacobian matrix:

$$\begin{pmatrix} \dfrac{s}{\alpha+1} h'_{\alpha}(s) & \dfrac{s}{\alpha+1} \\ \dfrac{\beta+1}{s} - h'_{\beta}(s) & 0 \end{pmatrix}, \tag{B.11}$$

where $h_{\mu}(x)$, which is introduced to simplify the notation, is given by:

$$h_{\mu}(x) = \int_0^x t^{\mu} f(t) \, dt \Big/ \int_0^s t^{\mu} f(t) \, dt. \tag{B.12}$$

To study the stability we look at the eigen-values of the Jacobian in eq. (B.11). The eigen-values satisfy the characteristic equation:

$$\lambda^2 - \text{Tr } \lambda + D = 0, \tag{B.13}$$

where

$$\text{Tr} = \frac{s^{\alpha+1}}{\alpha+1} \frac{f(s)}{\int_0^s t^{\alpha} f(t) \, dt} \tag{B.14}$$

and

$$D = \frac{\beta+1}{\alpha+1} \left(1 - \frac{s^{\beta+1}}{\beta+1} \frac{f(s)}{\int_0^s t^{\beta} f(t) \, dt} \right). \tag{B.15}$$

If the roots of eq. (B.13) would have magnitude less than 1, then (s, M) would be asymptotically stable. Conditions that must be satisfied guaranteeing the roots to have magnitude less than 1 are given by ref. [23]:

$$| \text{Tr} | < D + 1 < 2. \tag{B.16}$$

From expressions (B.14), (B.15) and using integration by parts as in eq. (B.10), it can be seen

that the conditions of eq. (B.16) are satisfied as long as $\beta < \alpha$ and that conditions (17) hold.

References

[1] M. Radermacher, J. Electron Microsc. Tech. 9 (1988) 359.

[2] H. Engelhardt, in: Methods in Microbiology, Vol. 20 (Academic Press, London, 1988) p. 357.

[3] M. Stewart and G. Vigers, Nature 319 (1986) 631.

[4] J. Dubochet, M. Adrian, J.-J. Chang, J.-C. Homo, J. Lepault, A.W. McDowall and P. Schultz, Q. Rev. Biophys. 21 (1988) 129.

[5] F.B. Booy, in: Membrane Fusion, Ed. J. Benz (CRC Press, Boca Raton, 1992), in press.

[6] T.S. Baker, J. Drak and M. Bina, Proc. Natl. Acad. Sci. (USA) 85 (1988) 422.

[7] S.D. Fuller, Cell 48 (1987) 923.

[8] B.V. Prasad, G.J. Wang, J.P. Clerx and W. Chiu, J. Mol. Biol. 199 (1988) 269.

[9] F.P. Booy, W.W. Newcomb, B.L. Trus, J.C. Brown, T.S. Baker and A.C. Steven, Cell 64 (1991) 1007.

[10] A.W. Agar, R.H. Alderson and D. Chescoe, Principles and Practice of Electron Microscope Operation (North-Holland, Amsterdam, 1974) p. 160.

[11] J.C.H. Spence, Experimental High Resolution Electron Microscopy (Oxford University Press, New York, 1980) p. 298.

[12] J.L. Carrascosa and A.C. Steven, Micron 9 (1978) 199.

[13] M. Unser, B.L. Trus and A.C. Steven, Ultramicroscopy 30 (1989) 299.

[14] D.B. Furlong, M.L. Nibert and B.N. Fields, J. Virol. 62 (1988) 246.

[15] M.V. Nermut, in: Animal Virus Structure, Eds. M.V. Nermut and A.C. Steven (Elsevier, Amsterdam, 1987) p. 373.

[16] F. Darcy and G. Devauchelle, in: Animal Virus Structure, Eds. M.V. Nermut and A.C. Steven (Elsevier, Amsterdam, 1987) p. 407.

[17] T.S. Baker, W.W. Newcomb, F.P. Booy, J.C. Brown and A.C. Steven, J. Virol. 64 (1990) 563.

[18] J.D. Schrag, B.V. Prasad, F.J. Rixon and W. Chiu, Cell 56 (1989) 651.

[19] N.H. Olson and T.S. Baker, Ultramicroscopy 30 (1989) 281.

[20] A. Klug and R.A. Crowther, Nature 238 (1972) 435.

[21] J. Frank and M. Van Heel, Ultramicroscopy 6 (1981) 187.

[22] M. Van Heel, Ultramicroscopy 13 (1984) 165.

[23] L. Edelstein-Keshet, in: Mathematical Models in Biology (Random House, New York, 1988) p. 57.

[24] G.A. Korn and T.M. Korn, in: Mathematical Handbook for Scientists and Engineers (McGraw-Hill, New York, 1968) p. 235.

[25] M.M.C. Bijholt, M.G. Van Heel and E.F.J. Van Bruggen, J. Mol. Biol. 161 (1982) 139.