A Measure for the Angle Between Projections Based on the Extent of Correlation Between Corresponding Central Sections

Ardan Patwardhan1*, Danielle Paul2, Hind A. AL-Khayat2 and Edward P. Morris2

1Department of Biological Sciences, Biochemistry Building Imperial College London London SW7 2AY, UK

2Biological Structure and Function Section, Biomedical Sciences Division, Imperial College London, London SW7 2AZ, UK

A pre-condition for the ab initio assignment of Euler angles to a set of projections from an asymmetric object is that at least three of the available projections correspond to rotations about different axes. For symmetric objects this condition may be relaxed. There are some applications of single-particle electron microscopy, such as the reconstruction of filamentous macromolecular assemblies, where all available projections more-or-less correspond to rotations about a common rotation axis making it difficult to satisfy this condition. Here, a method has been developed to overcome this problem, based on the fact that the correlation between two central sections of the Fourier transform of a compact object will not be limited to an infinitesimal central line but will have a finite extent, which is related to the angle between the corresponding projections. Projections from model filaments, with different degrees of rotational symmetry about the long axis, have been used to test the methodology. The results show that angle determination is robust down to signal-to-noise ratios as low as 2 and that, in general, the error decreases as the degree of symmetry increases. The method has been used to assign angles to a set of negatively stained muscle thick filament projections to obtain an initial 3D reconstruction. The main features of the projections are seen to be faithfully reproduced in the reprojections from the reconstruction. A real-space adaptation of this method is also discussed.

Introduction

The three-dimensional (3D) reconstruction of an object from its two-dimensional (2D) projections requires the knowledge of their relative angular orientations. In many applications, such as medical computerized tomography, the relative angular orientations can, with relative ease, be computed from the actual rotation(s) of the source, object and/or the detector.1 However, in single-particle electron microscopy (EM), the source and detector are fixed, and the projections of a large number of randomly oriented macromolecules are recorded. These projections tend to be very noisy, necessitating the use of multivariate statistical data compression2 and classification techniques3 to combine “like” views thereby obtaining projections with an enhanced signal-to-noise ratio (SNR). The relative angular orientations of these projections are not known initially and are commonly assigned using real4,5 or Fourier space methods6–8 that can be shown to be formally equivalent insofar as the basic principles are concerned.4 The Fourier space methods can be derived from the projection slice theorem,9 which states that the 2D Fourier transform (FT) of a 2D projection through a 3D density distribution is identical with a 2D central section, whose normal coincides with the direction of projection, of the 3D FT of the 3D density distribution.

Abbreviations used: EM, electron microscopy; SNR, signal-to-noise ratio; FT, Fourier transform; CTF, contrast transfer function; FSC, Fourier shell correlation; FWHM, full width half maximum; PDFs, probability density functions; SCE, sinogram correlation function; CRLP, correlation line profile.

E-mail address of the corresponding author:
a.patwardhan@imperial.ac.uk

© 2004 Elsevier Ltd. All rights reserved.

Keywords: 3D reconstruction; single particle analysis; electron microscopy; myosin; angular reconstitution
distribution. Thus, two central sections of the 3D FT of an asymmetric object will have a one-dimensional (1D) central line in common, which lies along the direction of the rotation-axis between the corresponding projections (note that henceforth the phrase “central section” automatically refers to the “central section of the 3D FT”). In the presence of symmetry, a central section may contain multiple common line pairs within itself, which can be used to orient it without the need for comparing it with other central sections. Real space methods of angle determination make use of the fact that two 2D projections will share at least one 1D line projection in the direction orthogonal to the rotation-axis, more in the presence of symmetry.

In the absence of symmetry, regardless of which method is used, three projections that do not share a common rotation axis are required to fix the orientations of the projections with respect to a reference system. This requirement may be relaxed to one projection if the object possesses 3-fold or higher order symmetry and if a projection is available whose normal is neither co-linear nor orthogonal to a 3-fold or higher order symmetry axis. Though this requirement may seem trivial to fulfill, and indeed is so for a number of macromolecular specimens, it is a limiting constraint for some, and an important class of such are the so-called “filamentous macromolecular assemblies”. These include extended polymeric systems, filaments made up of smaller globular subunits, or some combination of both. Examples include the thick and thin filaments of vertebrate muscle, which are responsible for muscle contraction and relaxation, as well as intermediate filaments, such as vimentin, which are critically involved in various cellular processes such as division.10 Due to the extended nature of these assemblies, they tend to lie within the plane of the specimen, restricting the availability of projections to those corresponding to rotations about the long-axis of the filament. A single projection is sufficient to complete a 3D reconstruction if helical symmetry can be assumed about the long-axis.11 While this, in many cases, is a very successful strategy, the method requires projection images corresponding to a complete asymmetric unit and is limited by any significant departure from helical symmetry. Thus, for example, the large asymmetric unit of the muscle thin filament and the long-range disorder which weakens the angular relationship between successive asymmetric units12 have so far precluded a complete 3D reconstruction by helical methods. Problems also arise with filaments characterized by systematic departures from helical symmetry of the type observed in muscle thick filaments.13,14 The type of limited angular distribution adopted by filamentous macromolecular assemblies also occurs in extended rod-like complexes such as the 26 S proteasome15 which may have little or no helical nature. A number of groups have adapted the basic single-particle scheme to allow for the study of particular filamentous systems16–23 but these do require a 3D model based on a structurally similar molecule and other a priori information to bootstrap the process and assign the initial set of angles to the projections. A drawback with this procedure is that even though the model might not be used in successive rounds of angular refinement, the final reconstruction could still be biased towards the original model. Additionally, in some cases, adequate models may simply not be available.

We have developed a method for estimating the angle between projections by exploiting the fact that as long as the volume of the object is finite, a certain degree of coherence exists between points in Fourier space:24–27 a coherence which, in general, decreases with the distance between the points and varies reciprocally with the volume of the object.26 The concept of “coherence volume”, the volume of a neighborhood in Fourier space within which there is high degree of coherence, is useful for illustrative purposes. The coherence volume is related to the Patterson’s interference function26 (also known as the G-function24,25). Figure 1(a) depicts the cross-section of two central sections, orthogonal to their rotation axis. Circles are used to illustrate coherence volumes centered on various highlighted points along the sections. If one considers pairs of points, (A,A’), (B,B’), (C,C’), etc. it is clear that the coherence decreases with the distance from the rotation axis. By comparing Figure 1(a)–(c), it can also be seen that coherence falls-off more quickly with radial distance as the angle between central sections increases. When the angle between central sections is between 90° and 180°, as in Figure 1(d), only the points close to the rotation axis have a high degree of coherence as we still compare (A,A’), (B,B’), etc. In this case, if one of the sections is mirrored with respect to the rotation axis before comparing points, one would obtain a profile similar for the case in Figure 1(b). Finally, a comparison of Figure 1(a) and (e) shows that rotation angles 0 and –0 will yield identical results, demonstrating the presence of hand ambiguity.

A detailed description of the method is presented. This is followed by model-data experiments that investigate the influence of symmetry, SNR, the contrast transfer function (CTF), and misalignments. These experiments by no means cover the full spectrum of parameters that may influence the methodology, or all the possible combinations thereof. In order to evaluate how well the method works on real data, we have assigned angles to two sets of class-averages of muscle thick filament projections. The first is a set obtained after first aligning the projections to a rectangular reference, followed by successive bouts of alignment against the average image of the full data set from the previous alignment round. Although the aim of this alignment procedure is to minimize translational and rotational misalignments of the individual projections, some misalignments will still remain. Additionally, in the Imagic image processing system (ImageScience GmbH) used here, the raw images are classified into an arbitrary number of
classes specified by the user. When the SNR of the projection images is high, it is possible to classify the projections so that there are only a few projections in each class. Conversely, when it is poor it is often necessary to have fewer but larger classes. Some degree of variation in the orientations of the projections constituting a class is therefore to be expected. Although the user will normally only use a subset of all the class-averages, deemed as being "good" based on a number of criteria (such as their visual sharpness) for further processing, the selected class-averages will still suffer from these problems though perhaps to a lesser extent than the ones excluded and if the new method is to be used for assigning initial angles to such class-averages, it must be robust with respect to such problems. One problem in assessing angle-assignment to the initial class-averages is that we do not have a priori knowledge of what the correct angles should be. We generate a 3D reconstruction from the initial class-averages based on the angles assigned, and then reproject the 3D reconstruction in the directions corresponding to the class averages (forward projections) and compare them with the class-averages visually. A good correspondence would suggest that a class-average was relatively homogeneous and that the angle assignment precise. However, this comparison is in itself inadequate to evaluate whether the method has been successfully applied, as it will in general be difficult to assess whether a poor correspondence is due to the quality of the class-averages, the angle assignment or both.

In order to make a direct assessment of the quality of angles assigned, we have used a second set of class averages, based on the same projection images as the first set, but which have undergone several rounds of model-based (starting from the 3D structure generated from the initial class-averages) alignment. Each round involves (a) generating forward projections from the 3D structure obtained in the previous round, (b) using these forward projections as reference...
images in realigning the raw projection images, (c) creating new class-averages from the latest alignment of the projection images, (d) assigning angles to the class-averages using angular reconstitution with forward projections from the previous round as references, and finally (e) generating a new 3D structure from the class-averages. That it is possible to use angular reconstitution in the refinement is due to the fact that once we have an initial 3D, we can project it out in any arbitrary direction, i.e. we are not limited to directions orthogonal to the long-axis of the filament. These steps can be iterated until the 3D structure is found to change little compared with previous iterations, either visually, by inspecting the correspondence between the forward projections and corresponding class-averages or by some resolution measure such as the Fourier shell correlation (FSC). If the iterative refinement procedure has worked, it is reasonable to expect that the final class-averages are of a better quality than the initial ones. Unlike in the case for the initial class-averages, here we do have angles to compare with, namely those assigned to the class-averages using angular reconstitution, and a direct assessment can therefore be made.

Theory

A measure for the correlation between central sections

We will here, for the sake of illustrating various aspects of the theory, make use of projections from model filaments as described in Model-based simulations. We begin by making the assumption that for a set of projections, the orientation of the rotation axis between any two is known. For filamentous proteins, the rotation axis would typically lie along the long axis of the filament, and for more globular proteins the orientations could be determined using angular reconstitution or common lines techniques. In order to determine the radial extent of the correlation between two central sections, the corresponding projections are first rotated so that the rotation axis in each projection is aligned with a predefined direction (Figure 2). They are then Fourier transformed in order to obtain the corresponding central sections, which can be represented by:

\[ F_1(u, v) = |F(u, v)| \exp[i \phi_1(u, v)] \quad \text{and} \quad F_2(u, v) = |F(u, v)| \exp[i \phi_2(u, v)] \]

respectively, where \(v\) is a pixel coordinate along the rotation axis, and \(u\) is a pixel coordinate orthogonal to it. Conjugate multiplication between the central sections then yields:

\[ F_1(u, v) F_2^*(u, v) = |F_1(u, v)| |F_2(u, v)| \exp[i(\phi_1(u, v) - \phi_2(u, v))] \]

and by taking the real-part of each pixel in the product image and normalizing it by its amplitude:

\[ \Re[F_1(u, v) F_2^*(u, v)] \]

we obtain an image where each pixel is directly related to the phase difference between the corresponding pixels in the two central sections. If the corresponding pixels in the two central sections are identical or differ only in amplitude, then the corresponding pixel in the final image will have the value of +1. A minimum value of -1 occurs if the phases are diametrically opposite. A 1D profile, hereinafter called the correlation line profile (CRLP), can then be calculated:

\[ \text{CRLP}(u) = \sum_v \frac{\Re[F_1(u, v) F_2^*(u, v)]}{|F_1(u, v)||F_2(u, v)|} = \sum_v \cos(\phi_1(u, v) - \phi_2(u, v)) \]

The summand is set to zero whenever the amplitudes of \(F_1\) and \(F_2\) are less than a predefined constant, in order to avoid division by zero. For a given pixel \(u\), the CRLP is a measure of the average phase difference for all \(v\): the smaller the average phase difference, the higher the value of the CRLP, the more random the phase difference, the closer the CRLP is to zero. Due to the presence of Hermitian symmetry in the central sections, CRLP will be an even function. It should be noted that the CRLP bears a resemblance to other measures that have been employed in EM to characterize the similarity between images, such as the Fourier ring correlation, the main differences being the normalization and the fact that summation is performed along lines rather than around rings.

Figure 3(a) depicts a set of CRLPs that are based on pairs of projections where the direction of the first projection is unchanged but that of the second is varied. The extent of correlation between central sections, as seen by the width of the peak, decreases substantially as the angle is increased to 90° but changes very little as the angle is increased beyond that. In order to characterize these curves using a single parameter, we have chosen to use the full width half maximum (FWHM) of the peak. However, FWHM estimates based directly on these curves could be inaccurate as they do not necessarily decrease monotonously from the peak value. After some experimentation, we found that Gaussian profiles fit quite well to the data (Figure 3), and hence we decided to use the FWHM of a Gaussian profile of the form:

\[ \text{CRLP}(u) = y_0 \exp\left[-4 \ln 2 \left(\frac{u - u_{\text{mid}}}{\text{FWHM}}\right)^2\right] \]

where the parameters \(y_0\) and FWHM are fitted to the data using least squares minimization. This is of course not the only way to characterize the curves and one could instead, for instance, use the total area under the curve normalized in some way. The
FWHM of the fitted Gaussian, however, did work satisfactorily under a broad range of circumstances. The variation of the FWHM with the angle between projections is shown in Figure 4(a). As has been mentioned, there is a broad range of angles over which the FWHM changes very little and this limits the usefulness of the FWHM as a direct measure of the angle between projections. However, we always have the possibility in software to create the mirror (with respect to the rotation axis) of a projection and as was shown in Introduction using Figure 1(d), the extent of correlation between two projections is likely to vary reciprocally with the extent of correlation when one of the two projections is mirrored. Figure 3(b) shows CRLPs where the first projection within the pairs is the...
Figure 3. Variation of the CRLP with the angle between projections. All CRLPs have been normalized in order to make the graphs easier to interpret. In all four cases shown, the angle of the first projection is kept constant, whereas that of the second is varied. (a) and (b) show CRLPs for a noiseless model filament with $C_1$ symmetry, the difference being that in (b) the first projection is mirrored with respect to the rotation axis before calculating the CRLP. (c) and (d) correspond to (a) and (b), respectively, except that these are based on a $C_1$ SNR=10 model filament. In all cases, the width of the central lobe of the CRLP varies with the angle; although there are broad ranges of angles overall the variation is small. The variation in (b) and (d) is approximately reciprocal to that in (a) and (c), respectively, a feature that is utilized in the definition of DFWHM.
same as in Figure 3(a), but has been mirrored with respect to the rotation axis prior to calculating the CRLP. The variation of the CRLPs with the angle between projections (prior to mirroring) can be seen to be approximately reciprocal to that in Figure 3(a). This reciprocity is even more apparent in Figure 4(a) where the FWHMs of CRLPs calculated using the same first projection but varying the direction of the second and the FWHM of corresponding curves where the first projection has been mirrored (FWHMM) are plotted versus the angle between projections. These curves are approximately mirror symmetric with respect to 0° due to the aforementioned hand ambiguity. Figure 3(c) and (d) correspond to Figure 3(a) and (b), respectively, except for the fact that these are based on noisy projections. Whereas noise influences the magnitude of the variation of the CRLP with projection angle, this variation is still present as is the reciprocity between FWHM and FWHMM. Given the reciprocal nature of the FWHM and the FWHMM, a single measure that is sensitive over the full asymmetric angle range can be obtained by taking the difference between them, i.e.:

\[ \text{DFWHM} = \text{FWHM} - \text{FWHMM} \]
DFWHM = FWHM - FWHMM  \hspace{1cm} (6)

which has a much smaller region of limited variation centered around 90° and is therefore likely to be a much more useful measure of the angle between projections. Figure 4(b) and (c) shows the variation of these measures in cases where the filaments possess 2 and 3-fold rotational symmetry about the filament (rotational) axis. The range of angles in either case is restricted to span an asymmetric unit. As with the \( C_2 \) model filament, these curves are mirror symmetric with respect to the centre of the angle range due to hand ambiguity. In the case of \( C_2 \) symmetry, a projection and its mirror are identical and therefore, the FWHM and FWHMM curves are identical and the DFWHM curve is approximately equal to zero. The same applies for higher order \( C_n \) symmetries where \( n \) is even. Thus, if we wish to use the DFWHM measure, we are restricted to examining filaments that possess odd rotational symmetries about the filament axis, see Limitations and future perspectives.

Converting DFWHM to angles

In general, the exact mapping from a DFWHM measurement to an angle is not known but has to be estimated from the set of DFWHM measurements. To that end, we approximate the relationship between the angle between two projections, \( \theta \), and the DFWHM by a sigmoid function of the form:

\[
\theta = \left( \frac{180}{NSYM} \right) \frac{1}{1 + \exp(\lambda \text{DFWHM})}, \lambda > 0 \hspace{1cm} (7)
\]

where \( \lambda \) is a fitting parameter and NSYM is the level of symmetry about the rotational axis, restricted to odd numbers. A comparison of the \( C_1 \) and \( C_2 \) DFWHM curves in Figure 4 shows that the \( C_2 \) curve is much more linear than the \( C_1 \) curve. This variation in linearity can be accommodated for by the \( \lambda \) parameter. According to the mapping function, the DFWHM range is unconstrained and can approach infinity but an examination of the curves in Figure 4 shows that the DFWHM does not actually tend toward infinity at the limits of the half angle range but seems to taper off towards a limiting value. This is especially evident if one examines Figure 4(c) in the region around 60°. Despite this disparity between the model function and the actual curves, we chose to use it as it modeled the rest of the curve quite well and required the estimation of only one parameter.

If \( \theta \) and DFWHM are random variables, then their probability density functions (PDFs) are related by:

\[
\text{PDF}(\theta) = \text{PDF}(\text{DFWHM}) \frac{1}{1 + \exp(\lambda \text{DFWHM})} \hspace{1cm} (8)
\]

under the assumption that \( \theta(\text{DFWHM}) \) is a one-to-one mapping,\(^{30} \) which is true if we use the function in equation (7). In order to derive an expression for \( \text{PDF}(\text{DFWHM}) \), one needs to make some assumption about the distribution of the measured angles and here we assume that they are uniformly distributed, i.e.:

\[
\text{PDF}(\theta) = \frac{NSYM}{180}, \theta \in \left[ 0, \frac{180}{NSYM} \right] \hspace{1cm} (9)
\]

For a real data set, the angle distribution may not be uniform and the potential consequences of this assumption are examined in Negatively stained thick muscle filaments. By substituting equation (9) and the derivative of equation (7) into equation (8), we obtain:

\[
\text{PDF}(\text{DFWHM}) = \frac{\lambda \exp(\lambda \text{DFWHM})}{(1 + \exp(\lambda \text{DFWHM}))^2} \hspace{1cm} (10)
\]

The parameter \( \lambda \) can be estimated from the standard deviation of a set of DFWHM measurements \( \lambda = \pi/3\sigma \). It was however found that the standard deviation was quite sensitive to outliers, which is why the average deviation is used instead:

\[
\sigma_{av} = \int_{-\infty}^{\infty} |\text{DFWHM}| \text{PDF}(\text{DFWHM}) d\text{DFWHM} = \frac{2 \ln 2}{\lambda} \hspace{1cm} (11)
\]

Angles relative to an absolute reference system

The angles obtained in the previous section are the acute angles between pairs of central sections. The next step is to use these relative angles to assign angles to the central sections relative to an absolute reference system. We first consider the case of \( C_1 \) symmetry. The fixed coordinate system selected is such that the angle for the first central section, \( \theta_1 \), within this system is set to zero. The angle of the second section relative to the first, \( \theta_{12} \), is assumed positive thereby introducing a preference of hand for the sections. Every possible set of sign combinations for the remaining relative angles is then tried, and the final angles are calculated using the sign combination that yields the most consistent angles in the absolute reference system. The sines and cosines of the absolute angle for section \( i \), expressed as an average of measurements taken relative to all other sections, is given by:

\[
\cos(\theta_i) = \frac{1}{N-1} \left[ \cos(\theta_{i1}) + \sum_{j=2}^{i-1} \cos(\theta_{ij} + \theta_{ij}) + \sum_{j=i+1}^{N} \cos(\theta_{ij} - \theta_{ij}) \right] ,
\]

\[
\sin(\theta_i) = \frac{1}{N-1} \left[ \sin(\theta_{i1}) + \sum_{j=2}^{i-1} \sin(\theta_{ij} + \theta_{ij}) + \sum_{j=i+1}^{N} \sin(\theta_{ij} - \theta_{ij}) \right] \hspace{1cm} (12)
\]

where \( N \) is the total number of central sections. In order to avoid problems associated with the cyclic nature of angles, the sines and cosines of the angles
are used rather than the angles themselves. By expanding the sine and cosine terms in equation (12), the set of equations for \( i = 2, \ldots, N \) can be formulated as a linear equation system and solved for the sines and cosines of the set of absolute angles. An example of this, when \( N = 4 \), is shown:

\[
\begin{pmatrix}
1 & 0 & -\cos\theta_{23} & -\sin\theta_{23} \\
0 & 1 & -\cos\theta_{23} & -\sin\theta_{23} \\
-\cos\theta_{23} & \sin\theta_{23} & 3 & 3 \\
-\sin\theta_{23} & -\cos\theta_{23} & 1 & 0
\end{pmatrix}
\begin{pmatrix}
\frac{3}{3} \\
\frac{3}{3} \\
\frac{3}{3} \\
\frac{3}{3}
\end{pmatrix}
= 
\begin{pmatrix}
\cos\theta_{24} - \sin\theta_{24} \\
\sin\theta_{24} - \cos\theta_{24} \\
\frac{3}{3} \\
\frac{3}{3}
\end{pmatrix}
\]

\[
= \begin{pmatrix}
\cos\theta_2 & \sin\theta_2 \\
\cos\theta_3 & \sin\theta_3 \\
\frac{1}{3} & \frac{1}{3} \\
\frac{1}{3} & \frac{1}{3}
\end{pmatrix}
\]

(13)

This is a symmetric positive definite equation system that can be solved using the LAPACK routine SPOSV.\(^{31,32}\) In order to evaluate the quality of the absolute angles obtained for a given sign combination, the total average deviation of the sines and cosines is calculated:

\[
\varepsilon = \frac{1}{2(N-1)} 
\sum_{i=1}^{N} \left\{ \cos(\theta_{ij}) - \cos(\theta_j) \right\} + \sin(\theta_{ij}) - \sin(\theta_j) \right\}
+ \sum_{j=2}^{N-1} \left\{ \cos(\theta_j - \theta_{ij}) - \cos(\theta_j) \right\} + \sin(\theta_j - \theta_{ij}) - \sin(\theta_j) \right\}
+ \sum_{j=i+1}^{N} \left\{ \cos(\theta_j - \theta_{ij}) - \cos(\theta_j) \right\} + \sin(\theta_j - \theta_{ij}) - \sin(\theta_j) \right\}
\]

(14)

The final angles are based on the sign combination that yields the lowest error.

If the angular domain is cyclic with a period that is smaller than 360° as is the case with higher order symmetries, equation (12) cannot be used directly as there will be multiple absolute angles that are equally valid. Instead we scale the relative angles prior to using equation (12), i.e.:

\[
\theta_{ij} = \frac{180}{N\text{SYM}} \theta_{ij}
\]

and then de-scale the absolute angles obtained by solving the linear equation system:

\[
\theta_k = \frac{\text{NSYM}}{180} \theta_k
\]

A set of \( N \) sections will yield a total of \( M = N(N-1)/2 \) relative angles. For a set of \( N \) sections, there are \( 2^M \) sign-combinations possible given that \( \theta_{12} \) is assumed positive. Thus, whereas an angle determination for a set of seven sections is completed within a matter of seconds on a Compaq EV6 workstation running True64 Unix, eight sections require about an hour, and larger sets have not been tried. In cases where the initial set of projections exceeds eight, it may be necessary to partition the projections into subsets of seven or eight projections and after initial runs on each subgroup, perhaps rearrange and regroup the projections and perform further runs. A program that calculates the angles for a set of projections has been implemented and is described in Software Implementation.

**Results**

**Model-based simulations**

The model-based simulations are described in Model-based simulations and Figure 5. The results of the SNR simulations are shown in Figure 5(c) with the average absolute difference between the angle estimated and the expected angle plotted versus the SNR categories. In general, it can be seen that at high SNR, the average absolute angular differences in the \( C_7 \) case are smaller than in the \( C_3 \) case which in turn are smaller than the \( C_1 \). This is to some extent what is expected given that the angular range of the asymmetric unit decreases with increasing symmetry. In general, for SNR values greater than or equal to 2, the differences are fairly constant and increase rapidly beyond this point. It is unclear as to why the SNR = \( \infty \) case has a higher error than the SNR = 10 case for \( C_1 \) and \( C_3 \) symmetries. One speculation is that interpolation artifacts (when forming the 2D projections from the 3D model) play a role here and that the presence of a small amount of Gaussian noise ameliorates the effects of such artifacts. The data point corresponding to SNR = 0.5 for the \( C_1 \) case is missing as the data in this case were so noisy that the \( \lambda \) parameter estimation did not work.

The results of simulations investigating the influence of the CTF on the methodology are shown in Figure 5(d) with the average absolute angular difference plotted against the defocus values. Except for the measurements at 1000 nm
Figure 5. Model-based experiments. (a) A projection image taken from the C₁ projection set with different levels of noise added to it; SNR shown above each image. (b) A projection image from the C₁ SNR=2 projection set with CTFs of different defoci imposed upon it; defocus value shown above each image. (c) The results of simulations in which noise has been added to model projections. The average absolute angular difference between the angle obtained using CRLP and the expected angle is plotted versus the SNR categories. The standard deviations of the angular differences were in all cases roughly equal to the average values and are not shown for the sake of clarity. (d) The results of simulations in which the C₁ SNR=2 projection set has had various CTFs imposed upon it. The average absolute angular difference between the angle obtained and the expected angle is plotted versus the defocus values used in the imposed CTFs. The error bars depict the standard deviation of the angular differences. (e) The results of simulations in which the C₁ SNR=2 projections have had uniformly distributed in-plane rotational misalignments imposed upon them. (f) The results of simulations in which the C₁ SNR=2 projections have had uniformly distributed in-plane translational misalignments orthogonal to the filament axis imposed upon them. In both (e) and (f), the average absolute angular difference between the obtained and expected angles is plotted versus categories of misalignment ranges tested. Every point in the charts (c)–(f) is an average based on five sets of seven angle measurements, i.e. 35 measurements.
defocus, the results seem to suggest that the method is relatively unaffected by the imposition of the CTF. Upon closer inspection of the 1000 nm defocus measurement, it was found that the main reason for the larger average difference as well as standard deviation was one single angle measurement that was off by 154°.

Simulations also revealed that the angle determination was fairly robust with respect to a few degrees of rotational misalignment between projections (Figure 5(e)), as well as a few pixels of transverse misalignment (Figure 5(f)). Simulations of longitudinal misalignment were unsuccessful (arbitrarily large λ values obtained) in all but for the ±1 pixel shift case. The higher sensitivity of the method to shifts in the longitudinal direction compared with the transverse direction can be understood from the fact that a shift in real space corresponds to a linear-phase exponential in Fourier space and from equation (4) it can be seen that while such a factor will affect the method regardless of the propagation direction, the effects are likely to be exacerbated for the component that coincides with the direction of summation, i.e. the longitudinal direction.

Negatively stained thick muscle filaments

Images of negatively stained vertebrate muscle thick filaments were used in order to evaluate the procedure with real data. Vertebrate thick filaments are made up of myosin molecules packed together into bipolar filaments with a roughly cylindrical backbone and myosin heads arranged in a three-stranded quasi-helical array on the filament surface. The filament images used here had been analysed using single particle methodology. In this previous analysis, images corresponding to class averages of filament segments (see Negatively stained thick muscle filaments for details of microscopy and image processing) were assigned angles with reference to a helical reconstruction. The resulting reconstruction proved to be a sufficiently good approximation to allow further refinement to a final 3D structure (Figure 6(a)) with well defined myosin head structure which revealed characteristic perturbations from helical symmetry.

Angles obtained for the same class averages using CRLP gave rise to an initial 3D map (Figure 6(b)) which already agreed remarkably well with the previous analysis (Figure 6(a)). Further refinement using the initial 3D map as reference resulted in a 3D structure (Figure 6(c)) which was extremely close to the 3D map from the earlier study. A more quantitative assessment of the accuracy of angular assignment obtained with CRLP can be made by considering the angles assigned to individual class averages.

As the actual angular orientations of the initial class-average images were not known, it was not possible to calculate the angular differences as was done for the model-based simulations. However, since six of the images were repeatedly used for the

![Figure 6. Equivalent surface renderings of 3D reconstructions of the thick filament obtained using PyMOL (DeLano Scientific, San Carlos, CA, USA): (a) based on the final class-averages and angles assigned using angular reconstitution, (b) based on the initial class-averages and angles assigned using CRLP, (c) based on the final class-averages and angles assigned using CRLP and (d) based on the initial class-averages and the γ angle assigned randomly (α=0°, β=90°). All reconstructions have had C₃ symmetry imposed upon them. Renderings in (a)–(c) are based on maps thresholded at the σ = 1 level and that in (d) is based on a map thresholded at the σ = 0.5 level. The scale bar represents 20 nm.](image)
estimation of the angles of the remaining class-average images, standard deviations of the angles for five (note that one of the six images was assigned the angle of 0°) of these can be estimated. The average angles for these five images were \(-49°, -47°, 6°, 7°, \) and \(18°\), and the corresponding standard deviations were \(7°, 6°, 8°, 5°, \) and \(5°\).

A comparison between selected class-average images and corresponding reprojections is shown in Figure 7(a). Whereas there is good correspondence

Figure 7. Angle assignment to muscle thick filament data. (a) A representative set of class-averages whose Euler angles (shown on the images themselves) are spread over the \(C_3\) angle range and have been determined using CRLP is shown on the top row and the corresponding set of reprojections below it. (b) The initial 3D model obtained using the angles determined by CRLP was used to boot-strap an iterative refinement process and the final class-averages obtained are shown in the top row. The second row shows corresponding reprojections from a 3D model after assigning angles using angular reconstitution. The third row shows corresponding reprojections from a 3D model after assigning angles by CRLP. The class-averages selected in (b) are those whose angles (when determined by angular reconstitution) are close to the ones in (a).
as far as the general features are concerned for three of these pairs, this is not true for the \((0^\circ, 90^\circ, -19^\circ)\) and \((0^\circ, 90^\circ, 49^\circ)\) orientations. The spread of angles was uneven for the data set as a whole with three dominant clusters of angles centered at about \(\pm 50^\circ\) and 0°, indicating the presence of preferred angular orientations (Figure 8(a)). After iterative refinement, the correspondence between class-average images and reprojections is improved over the entire asymmetric range of angles (Figure 7(b)), both if the angles are assigned using angular reconstitution or by CRLP. When angles were assigned using angular reconstitution, many of the final projections, however, had \(\beta\) angles that deviated significantly from 90° (Figure 8(b)). When the \(\gamma\) angles obtained using angular reconstitution are compared with those obtained with CRLP (Figure 8(c)), there is a linear trend with a slope of \(\sim 1\) and the average absolute deviation of the CRLP angles from the trend line is \(\sim 6^\circ\). The fact that the slope is not closer to unity and that the deviation is higher than expected from the simulations may to some extent be explained by the fact that the CRLP assumes \(\beta\) to be 90°, whereas the angle assignment using angular reconstitution clearly shows that this is not the case for all the projections. When 3D reconstructions from the initial class-averages with CRLP angle assignment, from the final class-averages with angles assigned using angular reconstitution and from the final class-averages with angles assigned using CRLP are compared (Figure 6), the gross features of the filaments can be seen to be the same and it is mainly in the detailed definition of the myosin heads that they differ. However, the main difference is between the reconstruction based on the initial class-averages and those based on the final ones and the differences between the two final reconstructions are on a more subtle level.

Discussion

Model-based simulations

The model-data experiments have allowed us to gauge the effect of some important parameters on the CRLP in otherwise ideal circumstances. In general, we found that the method works best with images that have a SNR greater than or equal to 2. In practical terms, this means that one is restricted to using class-average images which are, at least for cryo data, made up of a few tens of raw images. Though this might restrict one to using a small subset of the available class-average images, this may be sufficient for obtaining an initial model, which is the primary aim of this study. In general,

![Figure 8.](image)

(a) The Euler angle distribution, as obtained using the CRLP method, of the full 36 member initial set of class-averages from which those used in Figure 7(a) have been selected. (b) The Euler angle distribution, as obtained using angular reconstitution, of the full 29 member final set of class-averages from which those used in Figure 7(b) have been selected. The radial coordinate in these plots represents \(\gamma\) and the angular coordinate \(\beta\). In both Figures three main clusters can be seen at approximately \(\gamma = \pm 50^\circ\) and 0°. (c) Comparison of \(\gamma\) angles for the final set of class-averages obtained using angular reconstitution and the CRLP method.
the angle determination was seen to improve with increase order of odd symmetry. To a certain extent, this is what is to be expected as the range of angles within an asymmetric unit decreases with symmetry order. We, however, found that for all symmetry cases, that below a certain SNR, the errors increased dramatically or the method was simply unable to cope with the noise. Due to the normalization performed in equation (4), the CRLP should not be sensitive to amplitude variations, at least in the absence of noise. Such amplitude variations will occur due to the CTF of the microscope and in general one could expect a variation of the CTF between individual raw images, necessitating some form of CTF correction in order to preserve high-resolution information when the data are combined. The CTF experiments that were performed were very simple in that they assumed that the CTF was the same for all the projections, but they should be sufficient to assess whether the methodology can cope with such amplitude variations and, as was shown, the method did cope well over a broad range of defocus values.

The misalignment simulations showed as expected that the errors increased with misalignment but that a few degrees of in-plane rotational misalignment as well as a few pixels of transverse misalignment were tolerable and CRLP was more sensitive to longitudinal misalignments. Initial alignments are likely to be performed using model reference projections or by aligning the line projections of the projection images. It is difficult to quantify the level of misalignments that one is likely to encounter with real start-up projection images as this is likely to vary with the geometry of the filament as well as with the SNR of the projection images. Our general experience is, however, that in the critical longitudinal direction, shifts exceeding the tolerances of the method are not uncommon and may thus hinder its successful application.

**Negatively stained thick muscle filaments**

The aim of the experiment with muscle thick filament data was to assess the performance of CRLP with real data. The degree of agreement between the initial 3D map (Figure 6(b)) and the refined structures (Figure 6(a) and (c)) clearly demonstrates the utility of the current approach. Even at the initial stage of the analysis the majority of reprojections agreed quite well with the class averages (Figure 7(a)), although in some orientations the correspondence was less good. The correspondence was significantly improved, however, after successive rounds of refinement (Figure 7(c)). The fact that many of the final class-average images were found to have β angles that were off 90° by more than a couple of degrees may explain the higher than expected (from simulations) errors for the CRLP angle determination. However, if CRLP is to be used for assigning initial angles then it must be able to cope with such problems and this experiment thus shows that it indeed is capable of doing so.

One intrinsic assumption made in CRLP is that the relative angular orientations measured for a set of projections will be uniformly distributed. The determination of the λ parameter is based on this assumption. However, this assumption is clearly not valid for the thick filament data. In order to examine what impact, if any, this uneven distribution might have on angle determination, we obtained projections from the model C₃ filament (SNR = ∞) at angles corresponding to those determined for the initial dataset and assigned angles to them using CRLP. The average absolute angular deviation was ∼3°, which was only slightly higher than that obtained for a uniform distribution (see Figure 5(c)).

These results show that the CRLP method is capable of assigning angles to the vertebrate thick filament class averages reasonably faithfully, certainly sufficiently well to use as a reference for subsequent refinement. However, it is possible that other approaches might lead to 3D structures of sufficient quality to be used as references for refinement. To investigate this we calculated 3D maps using C₃ symmetry and assigning random angles to the class averages (Figure 6(d)). In this case the agreement of the 3D map with the refined maps (Figure 6(a) and (c)) was substantially worse than that obtained with CRLP. Nevertheless, each of these 3D maps, filtered to remove spurious high frequencies, proved an adequate reference for refinement. The observation that three independent reference structures each converged to effectively the same 3D structure lends considerable confidence to the original analysis. Of the initial reference structures, that obtained by CRLP was considerably closer to the final refined structure. In this case it appears that the more divergent references can also be successfully used for refinement. This probably arises from the combination of C₃ symmetry and strong low resolution features of the vertebrate thick filament. Nevertheless, it is generally desirable to use as accurate a reference as possible to avoid the risk of refinement leading to the wrong structure.

**A real space variant of CRLP**

Given the inherent similarities between the common lines technique and angular reconstruction, it is pertinent to inquire whether a real space variant of CRLP would be possible. In angular reconstruction, 1D line projections from two projections, taken over the full range of in-plane rotations, are compared with each other to form a sinogram correlation function (SCF) (Figure 9(a)–(e)). In essence, this is a 2D function where each axis corresponds to in-plane rotations of one of the two projections and the coordinates of the maxima of the SCF yield the orientations orthogonal to the rotation axis in the two projection images. The
Fourier space equivalents of the 1D line projections are central (origin intersecting) lines lying within the planes of the corresponding central sections; along the rotation axis, the central lines overlap completely. Examining the correlation between pairs of central lines at equal rotations from the rotation axis is equivalent to comparing line projections in real space at equal rotations from the direction orthogonal to the rotation axis. In the SCF, these correspond to coordinates lying along the diagonal centered on the maxima. The anti-diagonal centered on the maxima corresponds to peaks in the SCFs and the coordinates of these peaks yield the orientations of the rotation axis in the two images used to make the SCF. There are two peaks in each of the SCFs shown despite the fact that the data has $C_1$ symmetry simply due to the fact that the SCF is ambiguous with respect to hand. The red lines demarcate diagonals centred on the peaks, and the blue lines the anti-diagonals. Whereas there is strong density along the diagonal in (d), i.e. when the angle between the projections is small, the density along the anti-diagonal is stronger in (e), which implies that there is greater correlation between the projection at $150^\circ$ and the mirror of the $0^\circ$ projection, i.e. the $180^\circ$ projection, rather than with the non-mirrored projection itself. Density profiles of the SCF along the diagonal and the anti-diagonal for a range of angles between the projections (shown in the legends) are shown in (f) and (g), respectively. As can be seen, the profile changes gradually with the angle between projections but is plagued by the same sort of asymptotic behaviour that the FWHM of the CRLP exhibits. However, the variations in (f) and (g) are more-or-less reciprocal, which means that a measure that combines both the information along the diagonal and that along the anti-diagonal may function in an analogous manner to the DFWHM in the CRLP case.

Figure 9. A real-space measure for the angle between projections. Three sinograms from the $C_1$ model SNR=10 projection set with $\alpha=0^\circ$, $\beta=90^\circ$, and $\gamma=0^\circ$, $30^\circ$, $150^\circ$, are shown in (a), (b) and (c), respectively. The SCF between (a) and (b) is shown in (d) and that between (a) and (c) in (e). The intersections between the red and blue lines correspond to peaks in the SCFs and the coordinates of these peaks yield the orientations of the rotation axis in the two images used to make the SCF.
to the case where one of the projections can be considered mirrored. The correlation between central lines will decrease as the rotation from the rotation axis increases and the rate at which it falls-off can be expected to depend on the angle between projections. We examined diagonal profiles centered on SCF maxima for pairs of projections at a range of angles, and as can be seen from Figure 9(f) and (g), there indeed is a variation in these profiles with the angle. Given the shape of these curves, the total area under the curve rather than the FWHM may be an appropriate measure. By taking the difference between such a measure for the diagonal and anti-diagonal profiles, one can obtain a measure similar to the DFWHM and use it in precisely the same way. However, we have as of yet neither implemented the real space based method, nor have we assessed the relative merits of the two methods.

**Limitations and future perspectives**

The application of CRLP requires that the symmetry of the system is known as is the case with other methods for \textit{ab initio} angular assignment of single particle images such as angular reconstitution and common lines. In conventional single particle analysis the rotational symmetry can frequently be deduced by inspection of eigen images calculated as part of the image classification procedure.\textsuperscript{33} In the case of filamentous systems the limited range of views available probably precludes this type of approach. The 3-fold rotational symmetry of the vertebrate myosin filament used as a test object here is well established.\textsuperscript{34} For other systems where such prior knowledge is not available, information from other sources may be used to estimate a range of possible symmetries. It may then be possible to identify the symmetry by systematically assigning angles for each possible symmetry and evaluating the resulting reconstructions.

In its present form CRLP is only capable of dealing with filaments possessing odd symmetry. Even symmetry implies the presence of mirror symmetry in the projections, which precludes the use of DFWHM. This is, however, not a limitation \textit{per se} if one can find a functional mapping between FWHM values and angles, whose parameters could easily be estimated from statistical measurements of the FWHM values. We tried a number of mapping functions, but none were as robust as the sigmoid function which has proven to be the same for all projections, (c) the threshold level to be used, (d) the \( \lambda \) parameter, which, if set to zero, will be estimated by the program and (e) an option to either determine only the relative angles (relative mode) between projections, or to determine both the relative angles and the angles relative to an absolute reference system (absolute mode). As output, the program provides an estimate of the measured \( \lambda \) parameter, a table of the FWHM measurements, plots of the CRLPs and when determining absolute angles, the angles for the projections that resulted in the lowest error. For a large set of projections, one would typically run the program in relative mode on the full set of projections in order to estimate the \( \lambda \) parameter. The program would then be run in absolute mode using subsets of projections but using the estimated \( \lambda \) parameter.

**Materials and Methods**

**Software implementation**

A program, corlipo, that takes a set of projections as the input and calculates the angles between them has been implemented in Fortran 90 within the framework of the Imagic image processing system. General input parameters include (a) the name of the file containing the projections, (b) orientation of the rotation axis which is assumed to be the same for all projections, (c) the threshold level to be used, (d) the \( \lambda \) parameter, which, if set to zero, will be estimated by the program and (e) an option to either determine only the relative angles (relative mode) between projections, or to determine both the relative angles and the angles relative to an absolute reference system (absolute mode). As output, the program provides an estimate of the measured \( \lambda \) parameter, a table of the FWHM measurements, plots of the CRLPs and when determining absolute angles, the angles for the projections that resulted in the lowest error. For a large set of projections, one would typically run the program in relative mode on the full set of projections in order to estimate the \( \lambda \) parameter. The program would then be run in absolute mode using subsets of projections but using the estimated \( \lambda \) parameter.

**Model-based simulations**

Initial tests of the method relied upon model-based projections with various levels of noise added to them. All image processing was done using Imagic. A cubic volume of 128\(^3\) voxels with densities following a random Gaussian distribution was first generated. The volume was then band-pass filtered such that the residual transmission below the low frequency cut-off, the low frequency cut-off and the high frequency cut-off were 0.01, 0.02 and 0.3, respectively, where frequencies have been specified as fractions of the Nyquist frequency as will be the case hereinafter unless stated otherwise. A cylindrical mask with a diameter that was 0.2 times the length of a side of the cubic volume was applied to the volume to form a cylindrical structure that would serve as the object in tests involving \( C_1 \) symmetry. \( C_3 \) and \( C_7 \) test objects were formed by applying the appropriate symmetrization operators on the \( C_1 \) test object. Thirty six projections with equiangular increments spanning the asymmetric unit ([–180°,180°] in the \( C_1 \) case, [–60°,60°] in the \( C_3 \) case and [–25.71°,25.71°] in the \( C_7 \) case) were...
generated from each of the three test objects. Gaussian noise was then added to these projections to yield the following SNR categories: \( \infty, 10, 2, 1 \) and 0.5, thus resulting in five projection sets for each symmetry class (Figure 5(a)). The program was then applied on each of the 15 projection sets in order to perform relative angle measurements and obtain estimates for \( \lambda \). Five subsets of seven projections each were extracted at random from each of the projections sets. The program was then applied to each of the subsets in order to perform absolute angle measurements, using the \( \lambda \) parameters that were estimated from the corresponding full projection sets.

In order to examine the influence of the contrast transfer function (CTF) on the methodology, CTFs were imposed on the \( C_i \) SNR = 2 projection set. The approach used is similar to that described\(^{38} \) where the CTF is imposed on the noisy image, rather than noise being added to a CTF-imposed noise-free image, and implicitly assumes that the main source of noise is electron statistics. More precisely, five projection sets were generated from this data set, where the defocus assumed the following values: 500, 1000, 2000, 3000, 5000 nm (Figure 5(b)). The pixel size was set to 0.2 nm/pixel, the spherical aberration coefficient, \( C_a \) to 2 mm, and the acceleration voltage to 200 kV. The frequencies of the first zeroes for these defocus values are 1/1.1, 1/1.6, 1/2.1, 1/2.9 and 1/3.6 nm\(^{-1} \), respectively. The projection sets were then low-pass filtered using a Gaussian filter with a FWHM of 1/0.8 nm\(^{-1} \) in order to simulate the fall-off of the CTF. The projection sets were then masked using a radial mask with a fractional radius of 0.8 and a soft fall-off of 0.04 to minimize edge artifacts. Note that no CTF correction was performed, nor can it be assumed that the SNR was still equal to 2 after the additional steps had been performed. As before, relative angle measurements were performed on full projection sets in order to estimate \( \lambda \) for each defocus case and absolute angle measurements were performed on five subsets of seven projections each chosen at random from each of the projection sets.

In order to study the influence of misalignments of the projections, three simulations were carried out, one where the projections were rotationally misaligned, one where they were translationally misaligned along the axis of the filament (longitudinal misalignments) and one where they were translationally misaligned orthogonal to the filament axis (transverse misalignments). These simulations used the \( C_i \) SNR = 2 projection set and the imposed misalignments followed uniform random distributions. For the rotational misalignment simulations, eight projection sets were generated with misalignments ranging from \( \pm 2^\circ \) to \( \pm 8^\circ \). Likewise, for each of the translational misalignment simulations, eight projection sets were generated with misalignments ranging from \( \pm 1 \) to \( \pm 8 \) pixels in the appropriate direction. As for previous simulations, \( \lambda \) was estimated separately for each of the projection sets. Five subsets of seven projections were then chosen at random from each of the projection sets and absolute angle determination performed on each of these subsets.

**Negatively stained thick muscle filaments**

Micrographs of negatively stained thick muscle filaments from goldfish (*Carassius auratus*) were kindly provided by R. W. Kensler (Department of Anatomy, University of Puerto Rico Medical School, Puerto Rico). The filament preparation procedure, the negative staining procedure and the electron microscopy have been described in detail\(^{13} \). In brief, uranyl acetate was used for the negative staining; the filaments imaged using a JEOL1200EX electron microscope (JEOL USA Electron Optics, Peabody, MA) at 20,000×, and the micrographs digitized using Leafscansher-45 at 10 µm/pixel. The pixel size at the specimen was found to be 0.545 nm after calibration. Filament segments were interactively selected and cut out to 200×200 pixel images that were then treated as independent “particles”. These images were centered, aligned and subjected to multivariate statistical analysis using Imagic in order to combine “like” views thereby obtaining class-averages with improved SNR\(^{14,23} \).

Our initial attempts to use corlipro on these class-average images, despite the fact that the SNR seemed to be relatively good, resembling best the SNR = 10 images from the simulations, failed. Some experimentation revealed that a major contributing factor was the strong negative stain densities flanking the filament on either side. Therefore, a density threshold, \( T \), was selected, by visually inspecting the class-average images and their density histograms, which separated the strong negative stain artifacts from the densities attributed to the filament itself, and the images thresholded such that all densities below this value were set equal to it. A constant density value, \( -T \), was then added to the class-average images to reduce the background offset to zero. To further improve the quality of these images, they were masked using a rectangular mask with a width of \( \sim 60 \) pixels that had been low-pass filtered with a Gaussian filter with a FWHM of 0.5 in order to soften its edges. Finally, a radial mask with a fractional radius of 0.7 and a fall-off of 0.01 was applied.

The program was first run on the full set of classes in the relative measurement mode in order to estimate \( \lambda \). The maximum half-width half-maximum among the CRLP curves was \( \sim 1/5 \) nm\(^{-1} \), which in terms of the fraction of the Nyquist frequency is similar to some of the cases simulated. CRLP curves generated by the program for the individual measurements between class-averages (1260 measurements in total) were inspected visually to select class-averages that were giving the most stable measurements with respect to the rest of class-average images. A few test-runs of the program in absolute measurement mode were then carried out in order to pick out a sub-set of six class-average images that were well distributed throughout the asymmetric unit. corlipro was then run in absolute measurement mode on sets of seven images, comprising the six reference images and each of the other images, one at a time, i.e. it was run 30 times.

These angles were used to obtain a 3D reconstruction and reprojections from it were compared with the class-average images. The reprojections were also used to boot-strap an iterative refinement procedure where both the alignment of the raw images as well as angle determination by angular reconstitution were repeated a number of times. For the final set of class-averages, angles were calculated using both angular reconstitution and CRLP.

**Acknowledgements**

We thank Professor John M. Squire for discussions and support. We acknowledge support from the British Heart Foundation: D.P. is supported by a studentship (# 13962) and H.A.A-L.K. by a grant (# PG/02/054/13776), both awarded to John M. Squire. E.P.M. has a University Award from the
Wellcome Trust (# 066418). A.P. acknowledges funding from the BBSRC (B13016).

References


Edited by Sir A. Klug

(Received 11 May 2004; received in revised form 17 September 2004; accepted 23 September 2004)