Corrim-based alignment for improved speed in single-particle image processing

B. Sander, M.M. Golas, and H. Stark*

Max Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Goettingen, Germany

Abstract

The technique of single-particle electron cryomicroscopy is currently making possible the 3D structure determination of large macromolecular complexes at constantly increasing levels of resolution. Work at resolution now attainable requires many thousands of individual images to be processed computationally. The most time-consuming step of the image-processing procedure is usually the iterative alignment of individual particle images against a set of reference images derived from a preliminary 3-D structure. We have developed an improved multireference alignment procedure based on interpolated cross-correlation images (corrims) that results in an ∼8-fold acceleration of the iterative alignment steps. These corrims can be used to restrict the number of image-alignment calculations by narrowing down the set of reference images. Another improvement in alignment speed has been achieved by optimising the software and its implementation on many parallel processors. This new corrim-based refinement has been found to work well with two different alignment algorithms, the commonly used "fast alignment by separate translational/rotational searches" and "exhaustive alignment by polar coordinates."

1. Introduction

Single-particle electron cryomicroscopy is increasingly recognised as being powerful enough for structure determination of large macromolecular complexes at very high resolution even when no internal symmetry is present in the particle (Henderson, 1995; Jensen, 2001; van Heel et al., 2000). However, the number of individual molecular images needed is large, and computational image processing is very time-consuming. It is still unclear how many images will be needed for structure determination at close to atomic resolution. Various theoretical estimates have been made of the number of images that might be needed to reach ∼3 Å resolution (Henderson, 1995; Jensen, 2001). By extrapolating from our current experience, we expect that up to several hundred thousand images (without internal symmetry)

may be needed for single-particle structure determination at molecular resolution.

At least five parameters (two translational and three rotational) are required for a complete description of the individual images, and these have to be known for each image before a 3-D reconstruction can be computed. The need to determine these parameters is a consequence of the random orientation of the individual imaged particles and the variable displacement of the centre of the particle when it is picked out of the micrograph field. In the commonly used projection-matching technique, all these parameters are determined and refined by several iterations of a multireference alignment procedure (Penczek et al., 1994; Radermacher et al., 1986). The reference images used are projections of a preliminary 3-D density map, evenly distributed over the unit sphere. Each raw image is shifted and rotated with respect to its best-matching reference image. In addition, the Euler angles of this best-matching reference are assigned to the aligned image. The aligned images are then used to compute a new 3-D

* Corresponding author. Fax: +49-551-201-1197.
 E-mail address: holger.stark@mpibpc.mpg.de (H. Stark).

reconstruction, from which a refined set of references can be calculated. It is to be expected that, in the course of several alignment iterations, the noise of the reference images will lessen, and improved accuracy of the alignment parameters will ultimately be obtained.

The multireference alignment of thousands, or even hundreds of thousands, of particles is the most time-consuming step in the image-processing procedure. Aside from large numbers of images, a fine angular spacing of the reference images may be needed for high-resolution work. A given multireference alignment using an angular spacing of 3° between the reference images may thus utilise, e.g., ~250 000 raw images (each measuring 128×128 pixels) that have to be aligned against a set of ~4500 reference images. On the Compaq Alpha EV6 500 MHz computer (which is still commonly used), equipped with four CPUs and a standard installation of the IMAGIC-5 software package (van Heel et al., 1996), a single iteration would take about two years. It is obvious from this estimate that increasing computational power must be combined with improvements in software if such computationally demanding tasks are to be made feasible on a reasonable time scale. An increase in computational efficiency has so far been achieved by the use of algorithms for parallel processors on so-called "computer farms" (van Heel et al., 2000). In this paper, we describe "corrim-based alignment," a new approach that greatly increases the speed of alignment. We also describe its implementation on a PC cluster consisting of up to 68 AMD processors.

Alignment algorithms based on correlation functions are employed to orient individual particle images with respect to reference images (Saxton and Frank, 1977). In the present study, two different algorithms for the alignment of single image-reference pairs were applied. One of them is that of alignment by separate translational/rotational searches. In this approach, which is also called direct alignment (Harauz et al., 1988; Steinkilberg and Schramm, 1980), the rotational parameter is determined using the rotational correlation function, whereas the translational parameter is extracted from a 2-D cross-correlation function (CCF) in a separate step. The translational and rotational steps are repeated alternately 2–5 times until the image's position becomes stable. Alignment by separate searches has been successfully used and, largely by virtue of its speed, has become the de facto standard alignment method implemented in many commonly available image-processing packages. In contrast to separate alignments, the so-called "exhaustive alignment" methods are considerably slower. These algorithms evaluate all possible positions for a given molecular image within a defined range, and this makes them computationally more costly than direct alignment. However, the exhaustive "alignment using resampling to polar coordinates" is investigated here because of its excellent

accuracy for images with low signal-to-noise ratio (SNR) (Joyeux and Penczek, 2002). This alignment requires a transformation to polar coordinates for each possible centre of the particle in the pixel frame and scans directly all possible translational parameters without the need for 2-D CCFs. For each position a rotational cross-correlation is required to determine the possible rotation angle of the image. The input image to be aligned is shifted with respect to the reference image exhibiting the highest rotational correlation coefficient. The rotational position is determined according to the one-dimensional coordinate of the peak in the rotational correlation function.

In both multireference alignment procedures described here, Pearson's real-space correlation coefficients are computed for all image-reference pairs after completion of each individual alignment operation. These correlation coefficients are required to compare the results of the alignments of a single image against all references. The highest correlation coefficient determines the best-matching reference for any given particle image. The resulting alignment parameters are used to assign the shifts and Euler angles to the particle image. In the course of several alignment iterations, it is obviously desirable to avoid unnecessary time-consuming alignment calculations that involve reference images exhibiting reproducibly low correlation coefficients. For any given individual particle image, this can be achieved by restricting the iterative alignments to those references with known high correlation coefficients only. Under realistic conditions, a refined alignment restricted to the best matching reference of any given particle image only, is not sufficient. Therefore, when performing a first multireference alignment, we record the correlation values of each molecular image to the entire set of references. The set of correlation values of any given raw image is used to calculate an interpolated cross-correlation image—a so-called *corrim*. In the subsequent corrim-based refinements, alignment operations have to be computed with respect to those references in regions of high correlation coefficients only. The fraction of references to be omitted is adjusted by a threshold correlation value specified by the user. The performance and accuracy of *corrim-based alignment* were tested for both of the alignment algorithm types named above.

2. Methods

In the extension of the multireference alignment (MRA) algorithm presented here, the calculated correlation coefficients between all pairs of reference images and raw images are stored in a normal iteration of the multireference alignment, and their values are then represented by a spline interpolation as a 2-D function
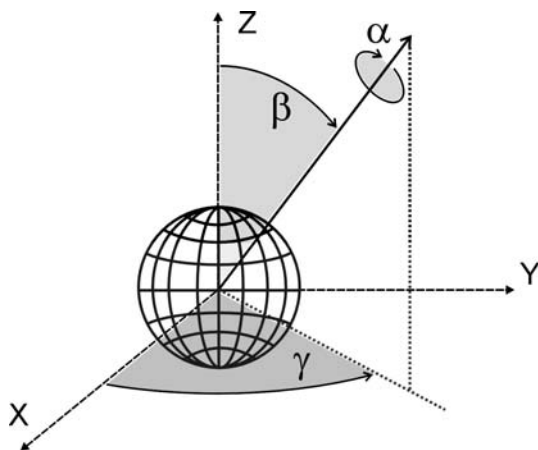
Fig. 1. Definition of the Euler angles for a right-handed, three-dimensional coordinate system, as used in the text. α defines the in-plane rotation of the projection. The position with respect to the "latitude" on an imaginary unit sphere is defined by β, while γ describes the position with respect to the "longitude."

of the Euler angles in a diagram, the ordinate of which represents the β angles and the abscissa the γ angles. The Euler angles were denoted according to (van Heel, 1987) (Fig. 1). The grey value at any point of such a cross-correlation image (corrim) can be interpreted as a measure of the probability with which a particle image can be assigned to a reference projection corresponding to the Euler angle defined by this point. The grey level thus determines whether or not the alignment to a specific reference image is computed in the subsequent MRAs. For this purpose a threshold value for the corrim is defined. Only when the threshold value has been exceeded is the alignment of the raw image and the corresponding reference image computed; otherwise, the programme moves on to the next reference image. The threshold value is defined in terms of the mean and standard deviation of the grey values in the corrim, leading to a normalisation of grey values. Prior to the simulations shown here, a sensible threshold parameter was determined empirically in multireference alignments of real data sets (the 70S ribosome (Stark et al., 2002) and the splicing factor SF3b (Golas et al., 2003)). It turned out that a conservative threshold of ∼1.2 times the standard deviation above the mean density of the corrim led to the computation of practically identical 3-D densities compared to a standard multireference alignment. By omission of alignment operations that are not needed, a significant reduction in processor time was achieved.

The multireference alignment presented here does not strongly depend on the kind of interpolation used in the calculation of the corrims. However, spline interpolations in this case are suitable to give the user the possibility to choose the size of the corrims freely and independently of the underlying angular distance of the reference images.

The spline condition used here for interpolation of correlation values, which fulfils the conditions of a monotonic first and a continuous second derivative, is given by (Press et al., 1997):

$$\frac{x_j - x_{j-1}}{6} y''_{j-1} + \frac{x_{j+1} - x_{j-1}}{3} y''_j + \frac{x_{j+1} - x_j}{6} y''_{j+1}$$
$$= \frac{y_{j+1} - y_j}{x_{j+1} - x_j} - \frac{y_j - y_{j-1}}{x_j - x_{j-1}}. \tag{1}$$

The 2-D interpolation in the corrims is achieved by performing two successive steps of 1-D interpolation that are orthogonal to one another. The first of these interpolations takes place between the projection angles at constant β and variable γ, that is, along a latitude line of the unit sphere. The second interpolation is then carried out in the longitudinal direction and always involves the correlation values measured for the two poles of the unit sphere as starting and finishing point, respectively. For n points (x_j, y_j) to be interpolated (where $j = 1, \ldots, n$), Eq. (1) leads to a system of $n - 2$ equations with n unknown variables, so that two additional conditions must be fulfilled for a unique solution to exist. For the longitudinal direction of interpolation, unambiguous definition of the system of equations is achieved by arbitrarily setting y''_1 and y''_n (the second derivatives of the spline curve measured at $\beta = 0°$ and $\beta = 180°$) to zero (the so-called *natural cubic spline*). In the case of the *cyclic* spline interpolation used in along a latitude line the two necessary additional conditions are given by taking account of the immediate neighbourhood of the first and last points of measurement:

$$\frac{x_1 - x_n}{6} y''_n + \frac{x_2 - x_n}{3} y''_1 + \frac{x_2 - x_1}{6} y''_2$$
$$= \frac{y_2 - y_1}{x_2 - x_1} - \frac{y_1 - y_n}{x_1 - x_n}, \tag{2}$$

$$\frac{x_n - x_{n-1}}{6} y''_{n-1} + \frac{x_1 - x_{n-1}}{3} y''_n + \frac{x_1 - x_n}{6} y''_1$$
$$= \frac{y_1 - y_n}{x_1 - x_n} - \frac{y_n - y_{n-1}}{x_n - x_{n-1}}. \tag{3}$$

Taking these additional conditions into account, the system of equations can be represented as a cyclic tridiagonal matrix with two additional boundary values (α, β), which can be solved efficiently by existing algorithms (Press et al., 1997):

$$\begin{pmatrix} b_1 & c_1 & 0 & \ldots & & & \beta \\ a_2 & b_2 & c_2 & \ldots & & & \\ & & \ldots & & & & \\ & & \ldots & a_{n-1} & b_{n-1} & c_{n-1} \\ \alpha & & \ldots & 0 & a_n & b_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_{n-1} \\ x_n \end{pmatrix}$$
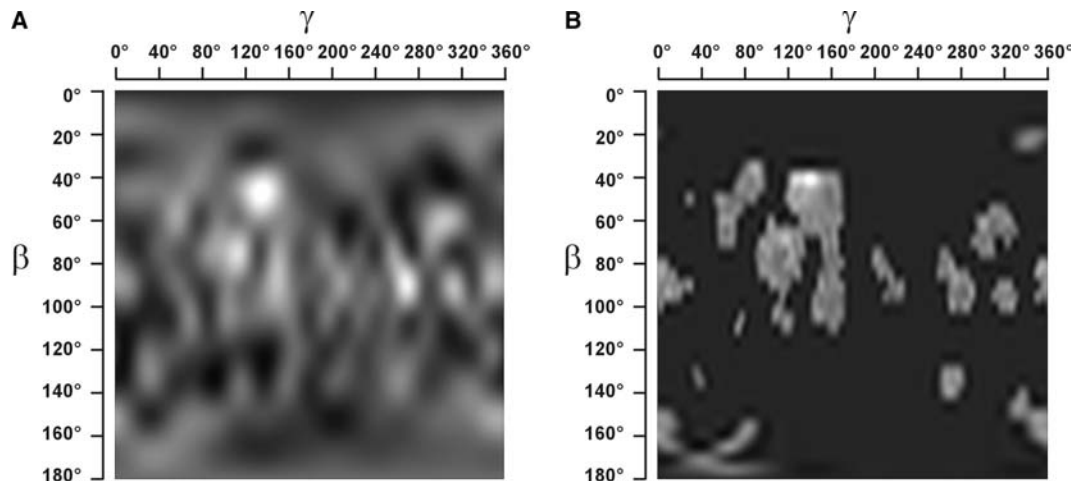$$= \begin{pmatrix} r_1 \\ r_2 \\ \ldots \\ r_{n-1} \\ r_n \end{pmatrix}. \tag{4}$$

Fig. 2. (A) Typical interpolated cross-correlation image ("corrim") after alignment of a particle image against a set of reference images with an average angular separation of 15°. The figure shows the progress of the expected correlation as a function of the Euler angles β and γ (see Fig. 1). (B) The same image after one iterative step of structure refinement, with (i) reference images separated on average by 4°, and (ii) account taken only of reference images with the projection angles that lay in regions with high correlation coefficients in the preceding alignment (light areas).

The result of the cyclic interpretation is a corrim, the left and right edges of which merge seemlessly into one another (Fig. 2). In the interpretation of the corrims, a distortion in the vicinity of the poles of the unit sphere can be discerned when the surface of the unit sphere is projected onto the plane. However, this distortion does not impede their function. Furthermore, the natural tendency of spline interpolations to oscillate, which can sometimes be seen in the vicinity of lighter regions, has not been found to have any disturbing influence upon the practical application in multireference alignment.

2.1. Preparation of the test data

The goal of the present study was to demonstrate the use of the method by applying it to a set of constructed test images that were as similar as possible to experimental electron-micrographic (EM) data. For this purpose, PDB data for the bacterial ribosomal subunits (50S, PDB entry 1FFK (Ban et al., 2000); 30S, entry 1FJG (Carter et al., 2000) and the [EF-Tu.aa-tRNA.GTP] ternary complex (PDB entry 1MJ1 (Stark et al., 2002)) were used to model a ribosome, and this was transformed into a 3-D density map by using the software package IMAGIC-5 (van Heel et al., 1996). To define the spectral signal decay, first, amplitudes in Fourier space were normalised by multiplication with the corresponding inverse rotational average, so that a signal strength independent of spatial frequency was obtained and the high spatial frequencies could be reduced by a Gaussian filter in an exactly defined manner.

The model measured $256 \times 256 \times 256$ pixels (1 pixel = 1.6 Å), whereas the alignment tests were carried out in (binned) 128×128-pixel windows. The pixel window was chosen such that the particle's largest diameter made up about two-thirds of the diameter of the window.

After Gaussian filtering of the amplitudes, the 3-D model had a signal that had decreased by a factor of 1/e at 12.8 Å and was set smoothly to zero at 4.5 Å (0.7 times the Nyquist frequency), as otherwise reconstructional artefacts of the 3-D algorithm used may appear in the region of very high spatial frequencies. To generate a model that includes the dark edges that are seen both around computed 3-D reconstructions and around 2-D class averages, a heavily low-pass-filtered version of the 3-D model was added to the unfiltered model; the tests described below were performed both with and without this additional component of low spatial frequencies, in order to investigate to what extent the results are influenced by their presence. The 3-D structure thus obtained was first projected with Euler angles that had been obtained by a random-number generator (*Prime Modulus Multiplicative Linear Congruential Generator* from the Compaq Fortran software-development package) to simulate projection images in the EM with random angular distribution. These projection images were then coarsened by averaging of 4-pixel fields, so that the 256×256-pixel images could be binned to a 128×128-pixel frame, simultaneously altering the critical Nyquist frequency to $1/6.4$ Å^{-1}. The projection images generated by the software IMAGIC-5 were then temporarily padded with zeros in a 256×256-pixel frame, so that there would be sufficient room for the broadening of points when convoluted with a CTF function based on a defocus of 10 000 Å, an assumed acceleration voltage of 200 kV and a spherical aberration of the electron lens of 2.0 mm.

Finally, to simulate defocus-corrected particle images, the images were subjected to defocus correction, so

that images with positive phases over the entire frequency interval and zero crossings corresponding to the defocus were received. The images were then rotated and displaced randomly, to simulate the random distribution of the particles in the EM field. The imposed rotation angles were in the range 0–360° and the imposed radial displacements were up to 12.8 pixels (i.e., ∼15% of the greatest particle diameter). To simulate various SNR values, a random-number generator was used to produce white Gaussian noise with a bell-shaped grey-value distribution, calculated as described by (Ross, 1997); the noise was added to the particle images. Taking into account the fact that most of the noise in the EM field is convoluted by a CTF, calculated noise images were overlaid with a CTF with the same parameters as for the test-image production. Thereafter, a low-pass filter was applied to the noise images, to simulate the—normally frequency-dependent—signal decay. The filter was chosen such that at 8 Å– i.e., later than for the projected images—a decrease to 1/e was reached, whereby at the same time the decrease in the spectral SNR expected in practice was also taken into account. Both the calculated particle images and the noise images were normalised in respect of their variance, so that defined SNR were obtained (Fig. 3); the SNR was defined as the quotient of the variance of the noise-free image and the variance of the additive noise (Frank and Al-Ali, 1975). By tests with experimental noise obtained from images of a carbon film in the EM, it was confirmed that there were no significant differences between the (more easily varied) calculated noise and the experimental noise. The size of the test-image data sets to be aligned lay between 500 und 5000 images and was chosen, on the basis of preliminary tests, so as to be large enough to yield highly reproducible results, with negligible scatter, when applied with various randomly-chosen parameter values and various noise values.

Reference images needed for all tests described in this paper were likewise calculated from the 3-D density maps generated by projection using the software IMAGIC-5 (see above). The angular separation between the reference data sets was chosen to be between 30° and 4°, so that reference image sets contained between 46 images (angular separation 30°) and ∼2500 images (separation 4°).

2.2. Conduct of the tests

The test images to be aligned, and the reference images, were generated as described above and a multireference alignment was carried out, so three Euler projection angles and two translational parameters were given to each simulated raw image by assigning it to the best-fitting reference image. The evaluation of these data was intended to answer the question of how the overall alignment error, resulting from the sum of individual errors, influenced the 3-D reconstruction. Back projections of single points (see below) in 3-D space were performed and the point-spread function resulting from the overall error was evaluated (by a method similar to that of (Jensen, 2001). Alignment errors can arise (i) through a deviation of the angle, assigned in the alignment, between the reference image and the aligned image, and (ii) through errors in the translation and rotation parameters found; the effects of these two factors influence each other. Therefore, a 3-D density map
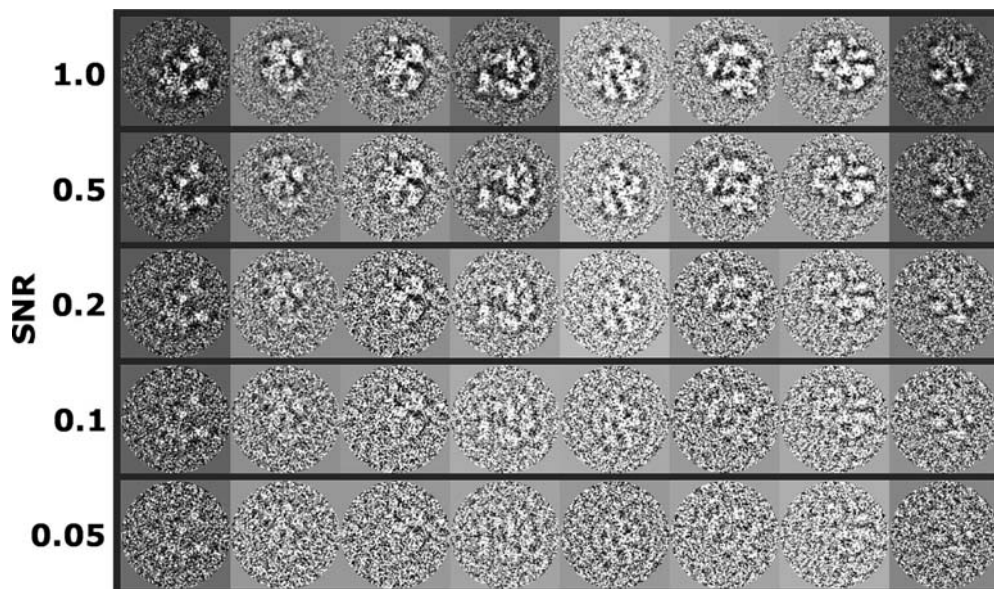

Fig. 3. Typical test images calculated on the basis of electron-microscope contrast theory as used in the multireference alignment. Each set of images was obtained for defined Euler angles and translational and rotational parameters (vertical groups) and is shown for SNR values that decrease from top to bottom. For details of the calculation, see text.

(128 × 128 × 128 pixel) was generated with a single image point that laid 40 pixels from the centre of the 3-D model (about 130 Å at the level of the constructed model—i.e., at the edge of the particle, where errors in the rotational parameters should have their greatest effect). This 3-D model was projected in all projection directions that had originally been assigned to the test-image data set. These "point projections" were then displaced/rotated by amounts corresponding to the respective differences between the imposed translational/rotational parameters and those found in the alignment. Furthermore, the Euler angles found for the test images in the multireference alignment were assigned to the corresponding "point projections." Afterwards, the set of displaced point projections was back-projected into 3-D space. The combination of all alignment errors caused a blurring of the point in the 3-D density.

To analyse this blurring, 3-D rotational averaging around the expected position was performed. By fitting the average grey values thus obtained to a decaying Gaussian function, the magnitude of the grey value at the origin voxel and the half-width of the signal decay were determined. The magnitude of the grey value was then compared with the highest possible grey value obtainable under ideal alignment conditions (usually found at SNRs greater than 0.3). It should be noted that the projection images of single points used for determining the envelope functions were set to be noise-free, so that any dependence of the envelope function on the number of particle images in the data set was avoided. In practice, the number of particle images that go into the calculation of a 3-D structure is naturally of great importance (Henderson, 1995). In consequence, the simulation experiments described here led to envelope functions that reflected the precision and significance of the results independently of the number of particle images used; that is, a larger number of particles in the simulation merely had the effect of reducing the scatter in the results.

3. Results and discussion

3.1. The influence of the corrim-based refinement on the computing time for multireference alignment

The speed with which an algorithm can be implemented on a given computer system depends upon many factors, and bottle-necks can arise at various points—for example, the transfer of data from the mass storage medium into the main memory of the computer (I/O time) or, especially with parallel processing, the transfer of data through the TCP/IP network. The option presented here for rapid refinement of 3-D models by multireference alignment makes very high demands on the computer, and it can only convey an advantage in terms of time if the system is able to transport the nec-

essary data—such as reference- and particle images—in and out of the computers used in a time less than the time required for processing.

Furthermore, it is theoretically possible that a refinement's advantage in terms of speed can shrink if the processor time required for the individual alignment of an image against a reference drops below a certain threshold. In such a case, a further reduction in the number of computing operations carried out would reduce the load on the processor but not the computing time. The attainment of a system-specific limit of this kind is, for the algorithms investigated, especially probable for alignment by separate searches, as this alignment is in itself much more compact than the exhaustive methods. To prevent the appearance of such effects with the use of corrims in alignment by separate searches, we adopted in the present work all available programming strategies to eliminate bottlenecks. So far, computational efficiency was enhanced mainly by establishing a memory buffer, which provided a high number of references preventing interruptions by I/O operations (van Heel et al., 2000). In addition, a similar buffer was set up for subsets of the particle images. These buffers led to more efficient data-package sizes to be transported across the TCP/IP-based network (100 Mbit) and decreased the frequency of transmissions, so that "data queues" were avoided. To perform the 2-D and 1-D Fourier transformations, the FFTW algorithm (Frigo and Johnson, 1998) was used, and routines for
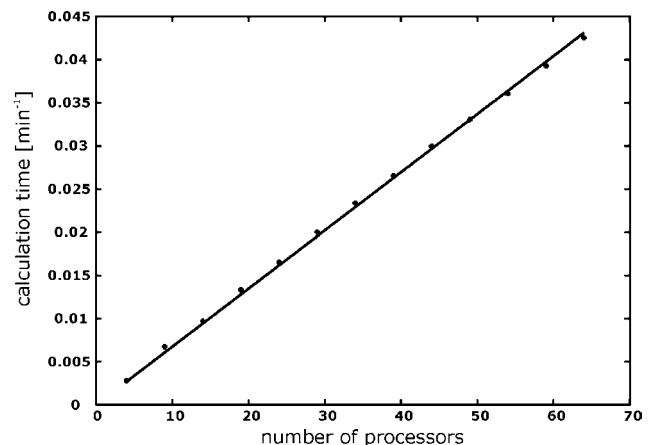

Fig. 4. Scalability of the refinement described here in the parallel version on 1–64 processors (AMD Athlon MP 2000+, Windows 2000 Prof., MPI-Software: NT-MPICH). Refinement of 20 000 test images measuring 128 × 128 pixels against 1650 reference images using corrims of 64 × 64 pixels and fast alignment by separate searches. Four of the processors of the type described performed the alignment in 5 h 57 min, while 64 processors required 23 min 30 s for the task. The reciprocal computing time in min^{-1} is plotted against the number of processors. In this example (which puts a high load on the network and I/O processes), the speed is almost proportional to the number of processors used. Only above 55 processors can a slight downward deviation be detected, which, however, appears not to be related to the increasing load upon the system, but rather to the initial distribution of reference images among the in-core buffers of the processors.

the translation/rotation of images and for the calculation of polar coordinates were optimised for the type of processor (AMD Athlon). For parallel computation on many processors, the software library NT-MPICH (http://www.lfbs.rwth-aachen.de/mp-mpich) was used. As shown in Fig. 4, the alignment by separate searches—which, when used with refinement by corrims, represents the most unfavourable case with respect to I/O and TCP/IP resources—proceeded at a rate that rose approximately linearly with the number of processors (up to 65 processors). For the refinements shown here, with a reasonable corrim size of 64×64 pixels, an approximately 8-fold increase in speed over conventional alignment was reproducibly obtained.

3.2. Characteristic appearance of corrims in dependence upon the SNR

Sharp peaks of correlation can be obtained with ideal (simulated) particle images and high signal-to-noise-ratios (Fig. 5). In real EM data sets, multiple correlation peaks may occur. This can have many reasons, including conformational inhomogeneity of the sample, local variation in quality of the exposure (Gao et al., 2002), differences in the amount of irradiation damage from image to image, low SNR values depending on the size of the imaged particles, characteristics of the surrounding medium (salt concentration, carbohydrates) and other factors. The characteristic change in
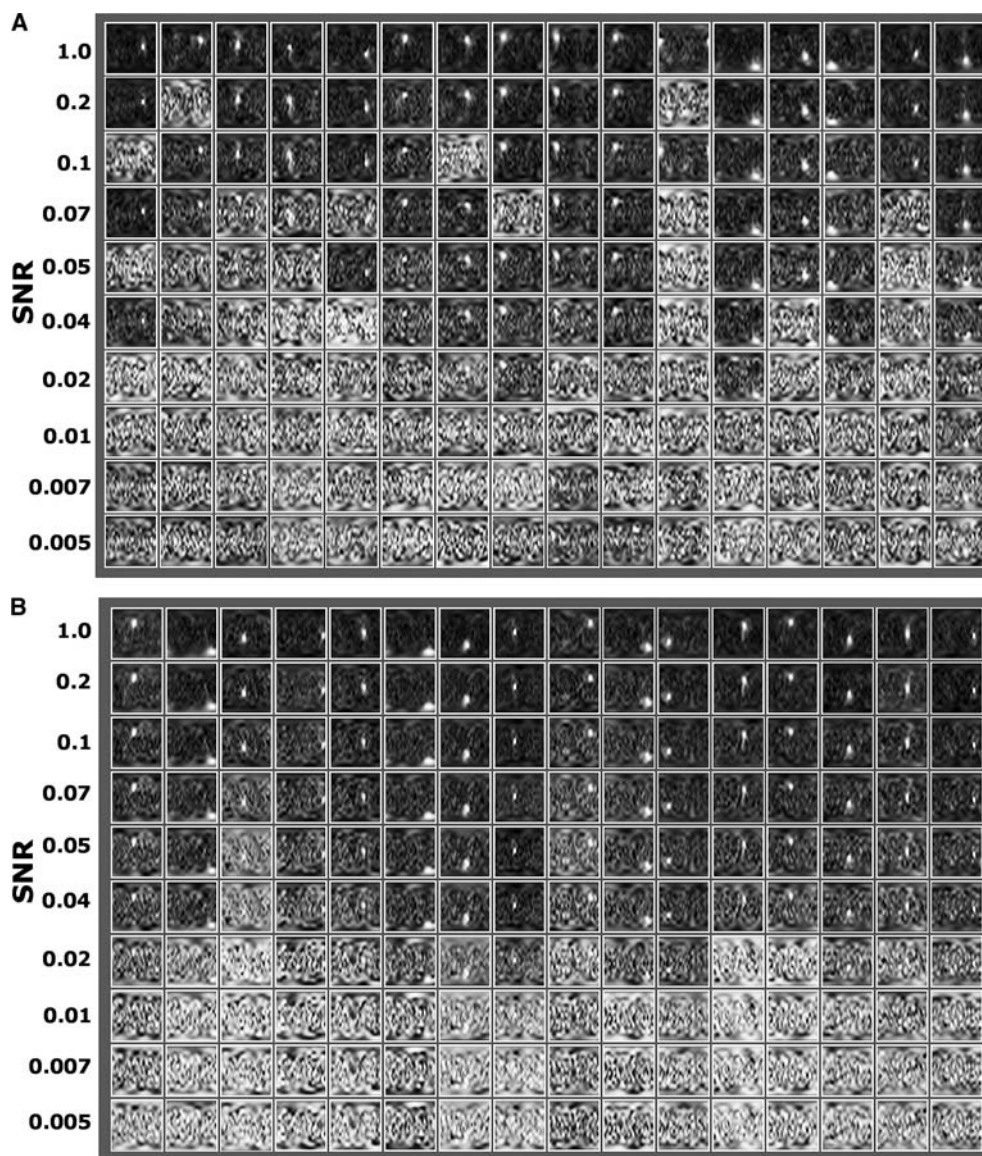

Fig. 5. (A) Appearance of the corrims as a function of SNR, as obtained in fast alignment by separate searches. A characteristic transition from unique to multiple correlation peaks can be observed with decreasing SNR. The individual SNR value where this transition occurs may differ from one image to another. At SNR values below 0.05, multiple peaks occur regularly. (B) The same as (A), but obtained in alignment by polar coordinates. It should be noted that the transition from an alignment that in principle is still correct to an alignment that is largely random takes place at lower SNR values than does the transition from unambiguous correlation peaks to multiple peaks (see text).
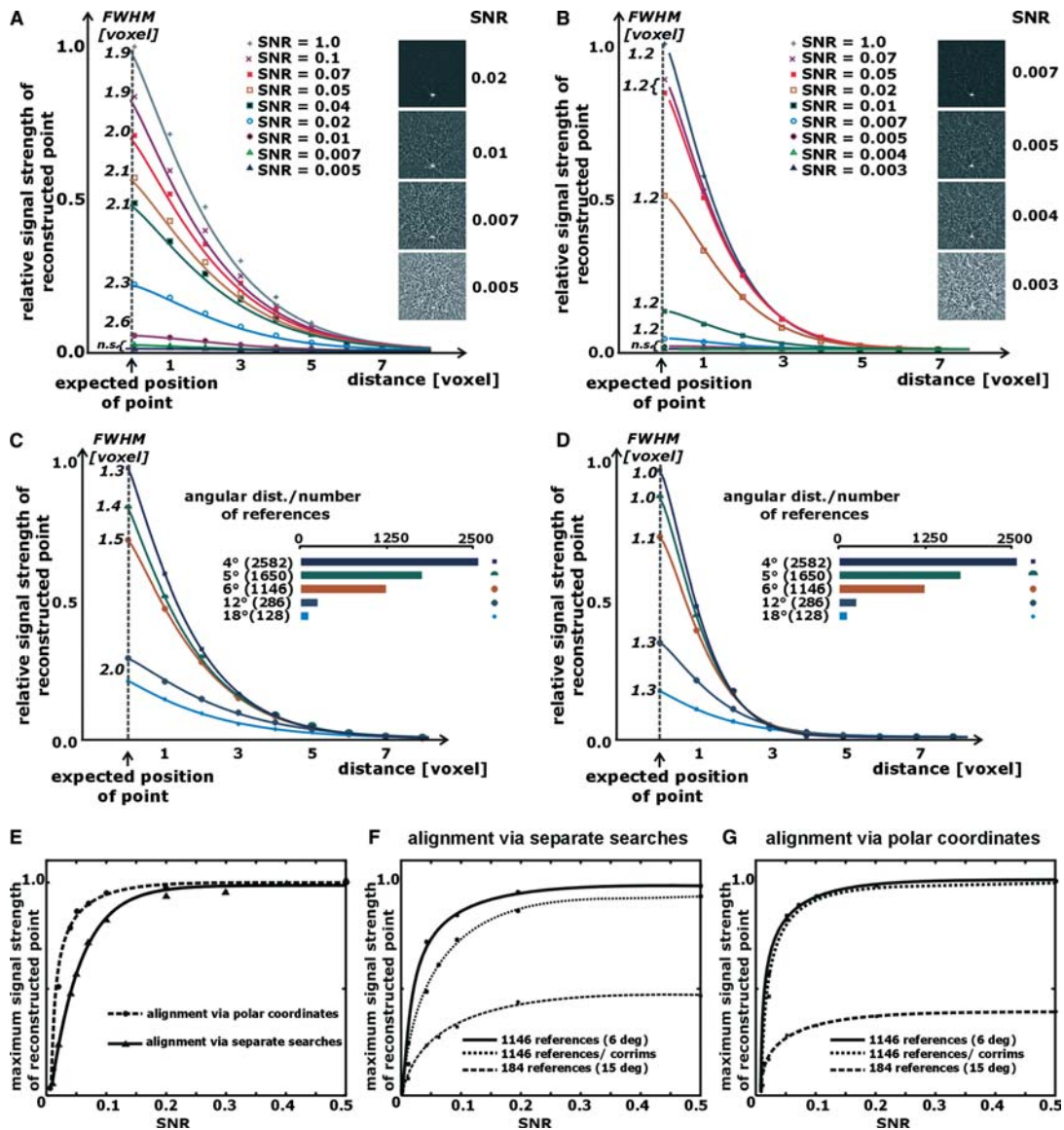
Fig. 6. Behaviour of the multireference alignment algorithm presented here, shown both with and without refinement of the test images in respect of (i) SNR and (ii) angular separation between the reference images. The overall error of the multireference alignment is made visible by superposing the measured individual errors upon projections of a single point, and the 3-D model of this individual point was reconstructed as described in the text. A. Influence of the SNR when alignment by separate searches is performed (5000 images, 286 reference images separated by 12°). The signal in the reconstruction of a single point obtained at SNR = 1 (see text) was arbitrarily assigned a signal strength of 1.0. The half-width in voxels, after rotational averaging about the expected point position, is shown in each case directly to the left of the maximum of the curve. The abrupt transition from a significant to a random-like result at an SNR of about 0.007 is illustrated by the 3-D reconstructions, shown in cross-section on the right. (B) As in (A), but with alignment performed by resampling to polar coordinates. The abrupt transition from a significant to a random result is seen here at an SNR of 0.004. (C) Influence of the size of the reference data set with alignment by separate searches (1500 images; SNR = 0.04). Between 128 and 2582 reference images were used in multireference alignment, corresponding to angular separations from 18° to 4°. Both the sharpness of the reconstructed image point—measured by the half-width of the rotationally averaged grey values—and, in particular, the signal strength increase with an increase in the number of reference images. Thus, an effect similar to that in (A) and (B) is seen at higher SNR values. The signal strength that was obtained in alignment against reference images with an angular separation of 4° was arbitrarily set to 1.0. (D) As in (C), but with alignment performed by resampling to polar coordinates (1000 images; SNR = 0.04). (E) The greatest signal strength obtained in (A) and (B) as a function of the SNR in alignment by separate searches and by polar coordinates. Both curves show plateaus over a wide range of SNR and collapse at low SNR values to zero; there is a region of low signal in which alignment by polar coordinates produces a significantly higher signal than that by separate searches. (F) Use of cross-correlation diagrams in alignment by separate searches. First, an alignment with 184 reference images (angular separation 15°) was performed. Cross-correlation images were calculated from the correlation coefficients obtained, and the signal strengths were plotted as in (E) (lower hatched line). These then served to restrict the set of reference images in a subsequent refinement step with 1146 reference images (angular separation 6°) so that an 8-fold acceleration of the computation was achieved in comparison with conventional alignment (signal strength: upper hatched line). The unbroken line shows the results of a conventional alignment in which the test images are oriented with respect to all 1146 reference images without use of the refinement option. The closeness of the two uppermost curves shows that the refinement option has worked well, giving a result only marginally worse than that of a conventional alignment with 1146 references (6°). The greatest signal strength that can be attained by conventional alignment against 1146 reference images (i.e., that at high SNR) was arbitrarily set to 1.0. (G) As in (F), but with alignment by polar coordinates. The comparison of the two uppermost curves shows that there is practically no difference between the conventional alignment and the refinement by corrims.

appearance of the corrims with decreasing SNR is shown in Fig. 5. In the alignment using resampling to polar coordinates, the transition from a single, clearly visible peak to multiple peaks occurs at lower SNR than in direct alignment. It can differ from one image to another. At SNR values below 0.04 (that is, for images in which the particles are scarcely visible; see Fig. 3) multiple peaks occur regularly. None the less, the simulation experiments shown here demonstrate that the alignment still functions well enough to produce meaningful results (see below). The threshold for the transition to random alignment results thus lies at significantly lower SNR values ($\sim$0.004–0.007) than does the threshold for the transition from unambiguous correlation peaks to multiple peaks.

3.3. Characterisation of the multireference alignment by signal strength and the half-width of a reconstructed voxel

As expected, the accuracy of the alignment parameter determination was dependent on the SNR. Interestingly, a decrease in SNR primarily reduced the maximum grey value of a reconstructed point, while the half-width did not depend strongly on the SNR (Figs. 6A and B). An increase in the number of reference images improved the signal strength similarly to an increase in SNR (Figs. 6C and D). Furthermore, a relatively narrow transition was observed from (in principle) correct functioning of the alignment to random-like results. This transition appears to be specific for the algorithm used; it occurs at lower SNR values, and more abruptly, for alignment by polar coordinates (0.004) than for alignment by separate searches (0.007-0.01, see Fig. 6E). As described above, test images with various proportions of low spatial frequencies were generated, thus simulating the situation frequently encountered in practice, e.g. as engendered by the use of different microscopes. It emerged that the result shown in Fig. 6E was in principle independent of the exact form of the spectral SNR curves. The appearance of a sharp boundary between conditions leading to correct and arbitrary alignment parameters confirms the results of a recent study in two dimensions (Joyeux and Penczek, 2002) and shows that its findings are equally applicable to multireference alignment and demonstrable in 3-D space.

In all the simulation experiments described here (Figs. 6A–G), the measured signal strength obtained was found to be the parameter that reacted most sensitively to changes in the alignment parameters such as SNR or the angular spacing of the reference images. It is clear that the signal strength, as measured, represents a complex function of the exactness of the alignment and the ratio between (in principle) correctly aligned and randomly misaligned images. In image-processing this signal strength would usually correspond to a subjective term such as "brilliance" or "intensity"; for the refinement of a 3-D model to the highest possible resolution it is more important that, with growing signal strength, an increasingly significant signal appears above the level of random noise, which leads ultimately to convergence of the structure model and, in turn, of the orientation parameters. The considerable increase in speed introduced by the corrim based alignment makes it possible to enhance the signal intensity by using larger data and reference sets, without incurring the disadvantage of longer computation time. Thus, a relatively small data set of approximately 180 reference images (angular separation $= 15°$) can be used to calculate the corrims and then, on the basis of this alignment, larger reference sets can be used. As Figs. 6F and G show, the signal strength measured over a wide range of SNR is comparable with a conventional alignment against a larger reference set (6°, $\sim$1100 reference images), whereby in the alignment with separate searches (using a threshold of 1.2 times the standard deviation above the mean density of the corrim) some 70–90% of the signal of a conventional alignment with 1100 reference images is reached, while with alignment using polar coordinates practically identical signal strengths (i.e., $\sim$100%) are reached over the entire range of SNR investigated.

4. Conclusions

The number of particle images needed to attain a desired resolution cannot be generally stated, as it is influenced by the symmetry and the molecular weight of the particle under investigation and also by the homogeneity, individual image quality and randomness of the angular distribution. The discrepancy between theoretically and experimentally attainable image quality increases the number of images from the theoretically required $\sim$10^4 (Henderson, 1995) to—according to our practical experience to date—some 10^{5–6}. Furthermore, the simulation experiments presented here indicate that the use of smaller angles between the reference images in the multireference alignment raises the quality of the calculated 3-D, and this improvement may be relevant for the success of the iterative refinement. The SNR of the individual particle images has been described as a decisive parameter for the convergence of cross-correlation-based methods to give the correct alignment (Sigworth, 1998). On the basis of simulated data, the present study indicates that higher numbers of reference images, i.e. smaller angular spacings, may have an influence on the convergence of the multireference alignment algorithm similar to that of high SNR values of the raw images.

The administration and processing of such large data sets have been made possible by the current rapid development in available hardware. The method presented here improves the efficient use of a large number of particle and reference images in multireference alignment

and, thanks to its highly parallel nature, can be efficiently performed in a "computer farm" environment.

In our experience, corrim-based alignment is also useful when the sampling interval (Å per pixel) of the particle images is reduced in the course of the structure refinement: a transition to a version of the particle images with a higher Nyquist frequency causes—through the enlargement of the image size—a decrease in the alignment speed in rough proportion to the square of the enlargement. The previously determined alignment information (contained in the corrims) can directly be used for the data set with enlarged pixel window and thus conveys a very large advantage in terms of speed. Furthermore, we observe an additional effect: it is known that with decreasing size of the sampling interval, the statistical significance of the correlation peak decreases (Frank, 1982). Accordingly, large sampling intervals result in more significant correlation peaks. Using corrims, the cross-correlation information determined from an alignment using large sampling intervals can be applied to subsequent rounds of the iterative refinement independent of the actual sampling interval. Thus, corrim-based alignment may have a favourable influence upon the quality of the final result when increasingly smaller sampling intervals are used.

At present, the question of how the information content of corrims is to be interpreted remains unanswered. It is generally accepted that the values of individual correlation coefficients in no way allow statements to be made about the quality or the correctness of an alignment. However, the information content of a family of correlation coefficients should be significantly greater. This may lead to an interpretation of the grey-value distribution in the corrims in terms of the qualitative assessment of an alignment. We are currently investigating this question.

Acknowledgments

References

Ban, N., Nissen, P., Hansen, J., Moore, P.B., Steitz, T.A., 2000. The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. Science 289 (5481), 905–920.

Carter, A.P., Clemons, W.M., Brodersen, D.E., Morgan-Warren, R.J., Wimberly, B.T., Ramakrishnan, V., 2000. Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics. Nature 407 (6802), 340–348.

Frank, J., 1982. New methods for averaging non-periodic objects and distorted crystals in biologic electron microscopy. Optik 63, 67–89.

Frank, J., Al-Ali, L., 1975. Signal-to-noise ratio of electron micrographs obtained by cross-correlation. Nature 256, 376–378.

Frigo, M., Johnson, S.G., 1998. FFTW: An adaptive software architecture for the FFT. Proc. IEEE Int. Conf. Acoust. Speech Sig. Process. 3, 1381–1384.

Gao, H., Spahn, C.M., Grassucci, R.A., Frank, J., 2002. An assay for local quality in cryo-electron micrographs of single particles. Ultramicroscopy 93 (2), 169–178.

Golas, M.M., Sander, B., Will, C.L., Luhrmann, R., Stark, H., 2003. Molecular architecture of the multiprotein splicing factor SF3b. Science 300 (5621), 980–984.

Harauz, G., Boekma, E., van Heel, M., 1988. Statistical image analysis of electron micrographs of ribosomal subunits. Methods Enzymol. 164, 35–49.

Henderson, R., 1995. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. Q. Rev. Biophys. 28 (2), 171–193.

Jensen, G.J., 2001. Alignment error envelopes for single particle analysis. J. Struct. Biol. 133 (2–3), 143–155.

Joyeux, L., Penczek, P.A., 2002. Efficiency of 2-D alignment methods. Ultramicroscopy 92 (2), 33–46.

Penczek, P., Grassucci, R.A., Frank, J., 1994. The ribosome at improved resolution: New techniques for merging and orientation refinement in 3-D cryoelectron microscopy of biological particles. Ultramicroscopy 53, 251–270.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1997. Numerical Recipes in Fortran 77. Cambridge University Press, Cambridge.

Radermacher, M., Frank, J., Mannella, C.A., 1986. A new 3-D reconstruction scheme applied to the 50S ribosomal subunit of E. coli. J. Microsc., 141.

Ross, S.M., 1997. A First Course in Probability, fifth ed. Prentice Hall.

Saxton, W.O., Frank, J., 1977. Motif detection in quantum noise-limited electron micrographs by cross-correlation. Ultramicroscopy 2, 219–227.

Sigworth, F.J., 1998. A maximum-likelihood approach to single-particle image refinement. J. Struct. Biol. 122 (3), 328–339.

Stark, H., Rodnina, M.V., Wieden, H.J., Zemlin, F., Wintermeyer, W., van Heel, M., 2002. Ribosome interactions of aminoacyl-tRNA and elongation factor Tu in the codon-recognition complex. Nat. Struct. Biol. 9 (11), 849–854.

Steinkilberg, M., Schramm, H.J., 1980. Eine verbesserte Drehkorre-lationsmethode für die Strukturbestimmung biologischer Makro-moleküle durch Mittelung elektronenmikroskopischer Bilder. Z. Physiol. Chem. 361, 1363–1369.

van Heel, M., 1987. Angular reconstitution: a posteriori assignment of projection directions for 3-D reconstruction. Ultramicroscopy 21, 111–124.

van Heel, M., Gowen, B., Matadeen, R., Orlova, E.V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M., Patwardhan, A., 2000. Single-particle electron cryo-microscopy: towards atomic resolution. Q. Rev. Biophys. 33 (4), 307–369.

van Heel, M., Harauz, G., Orlova, E.V., Schmidt, R., Schatz, M., 1996. A new generation of the IMAGIC image processing system. J. Struct. Biol. 116, 17–24.