cation second' approach. It is possible to reverse this order by using invariants with the supporting rationale that once approximately homogeneous classes of images were found, it should be easy to align them subsequently as within each class they will share the same motif.

A practical approach to reference-free alignment known as *alignment by classification* (Dube *et al.*, 1993) is based on the observation that for a very large data set and centred particles one can expect that although the in-plane rotation is arbitrary, there is a high chance that at least some of the similar images will be in the same rotational orientation. Therefore, in this approach the images are first (approximately) centred, then subjected to classification, and subsequently aligned.

In its simplest form, the *multireference alignment* belongs to the class of *supervised classification* methods: given a set of templates (*i.e.*, reference images; these can be selected unprocessed particle projections, or class averages that resulted from preceding analysis, or projections of a previously determined EM structure, or projections of an X-ray crystallographic structure), each of the images from the available data sets is compared (using a selected discrepancy measure) with all templates and assigned to the class represented by the most similar one. Equally often multireference alignment is understood as a form of *unsupervised classification*, more precisely *K*-means classification, even if the description is not formalized in terms of the latter. Given a number of initial 2D templates, the images are compared with all templates and assigned to the most similar one. New templates are calculated by averaging images assigned to their predecessors and the whole procedure is repeated until a stable solution is reached.

2.5.7.7. Initial determination of 3D structure using tilt experiments

The 2D analysis of projection images provides insight into the behaviour of the protein on the grid in terms of the structural consistency and the number and shape of projection images. In order to obtain 3D information, it is necessary to find geometrical relations between different observed 2D images. The most robust and historically the earliest approach is based on tilt experiments. By tilting the stage in the microscope and acquiring additional pictures of the same area of the grid it is possible to collect projection images of the same molecule with some of the required Eulerian angles determined accurately by the setting of the goniometer of the microscope.

In random conical tilt (RCT) reconstruction (Radermacher *et al.*, 1987), two micrographs of the same specimen area are collected: the first one is recorded at a tilt angle of ~50° while the second one is recorded at 0° (Fig. 2.5.7.3). If particles have preferred orientation on the support carbon film (or within the amorphous ice layer, if no carbon support is used), the projections of particles in the tilted micrographs form a conical tilt series. Since in-plane rotations of particles are random, the azimuthal angles of the projections of tilted particles are also randomly distributed; hence the name of the method. The untilted image is required for two reasons: (i) the particle projections from the untilted image are classified, thus a subset corresponding to possibly identical images can be selected ensuring that the projections originated from similar and similarly oriented structures; and (ii) the in-plane rotation angle found during alignment corresponds to the azimuthal angles in three dimensions (one of the three Eulerian angles needed). The second Eulerian angle, the tilt, is either taken from the microscope setting of the goniometer or calculated based on geometrical relations between tilted and untilted micrographs. The third Eulerian angle corresponds to the angle of the tilt axis of the microscope stage and is also calculated using the geometrical relations between two micrographs. In addition, it is necessary to centre the particle projections selected from tilted micrographs; although various correlation-based schemes have been proposed,

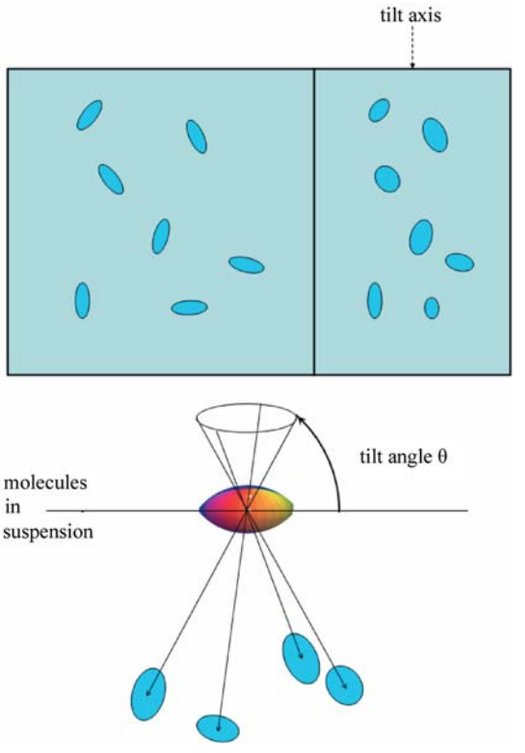

Fig. 2.5.7.3. Principle of random conical tilt reconstruction. A tilt pair of images of the same grid area is collected. By aligning the particle images in the untilted micrograph (left), the Eulerian angles of their counterparts in the tilted micrograph (right) are established. The particle images from the tilted micrograph are used for 3D reconstruction of the molecule (bottom). The set of projections form a cone in Fourier space; information within the cone remains undetermined.

the problem is difficult as the tilt data tend to be very noisy and have very low contrast.

Given three Eulerian angles and centred tilted projections, a 3D reconstruction is calculated. There are numerous advantages of the RCT method. (i) Assuming the sign of the tilt angle is read correctly (it can be confirmed by analysing the defocus gradient in the tilted micrographs), the method yields a correct hand of the structure. (ii) With the exception of the in-plane rotation of untilted projections, which can be found relatively easily using alignment procedures, the remaining parameters are determined by the experimental settings. Even if they are not extremely accurate, the possibility of a gross error is eliminated, which positively distinguishes the method from the *ab initio* computational approaches that use only untilted data. (iii) The computational analysis is entirely done using the untilted data, which have high contrast. (iv) The RCT method is often the only method of obtaining 3D information if the molecule has strongly preferential orientation and only one view is observed in untilted micrographs. The main disadvantage is that the conical projection series leaves a significant portion of the Fourier space undetermined. This follows from the central section theorem [equation (2.5.6.8) of Section 2.5.6]: as the tilt angle is less than 90°, the undetermined region can be thought to form a cone in three dimensions and is referred to as the missing cone. The problem can be overcome if the molecule has more than one preferred orientation. Subsets of particles that have similar untilted appearance (as determined by clustering) are processed independently and for each a separate 3D structure is calculated. If the preferred orientations are sufficiently different, *i.e.*, the orientations of the original particles in three dimensions are sufficiently different in terms of their angles with respect to the z axis, the 3D structures can be aligned and merged, all but eliminating the problem of the missing cone and yielding a robust, if resolution-limited, initial model of the molecule (Penczek *et al.*,