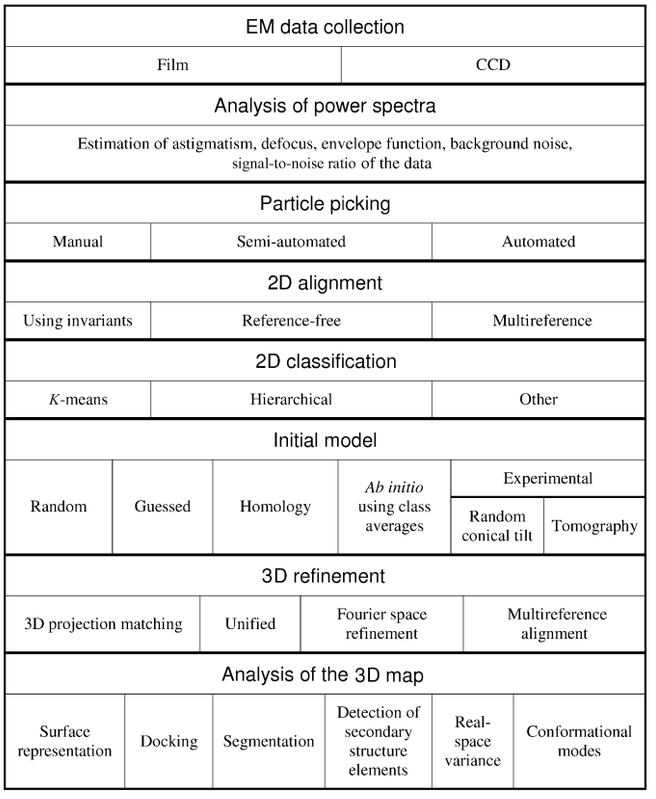

Fig. 2.5.7.1. Typical steps performed in a single-particle cryo-EM structure determination project.

between the orientation parameters $\psi_n, \theta_n, \varphi_n, t_{x_n}, t_{y_n}$ (n = 1, 2, . . . , N) and the 3D density f.

The parameters in (2.5.7.6) to be determined can be separated into two groups. (1) The orientation parameters $\psi_n, \theta_n, \varphi_n, t_{x_n}, t_{y_n}$ that have to be determined entirely by solving (2.5.7.6) and for which there are no initial guesses, and the structure f itself, for which we may or may not have an initial guess. The number of parameters in this group is very large: $n^3 + 5m$. Note that in single-particle reconstruction, the number of projection data m is far greater than the linear size of the data in pixels, *i.e.*, $m \gg n$. (2) Various parameters which we will broadly call the parameters of the image formation model (2.5.7.1)–(2.5.7.4): the defocus settings of the microscope Δf_n, the amplitude contrast ratio q and, if analytical forms of the envelope function E, the power spectrum of the background noise M, or the structure F are adopted, the parameters of these equations. Some of the parameters in the second group are usually known very accurately or can be estimated from micrograph data before one attempts to solve (2.5.7.6) (see Section 2.5.7.4), but they can also be refined during the structure determination process [for the method for correcting the defocus settings, see Mouche *et al.* (2001)].

Owing to the very large number of parameters in (2.5.7.6) and the nonlinearities present, one almost never attempts to solve the problem directly. Instead, structure determination using the single-particle technique involves several steps. (i) The macromolecular complex is prepared with a purity of at least 90%. (ii) The sample is flash-frozen in liquid ethane. Alternatively, cryo-negative stain techniques or traditional negative stain methods can be used. (iii) Pictures of the macromolecular complexes are taken. (iv) Exhaustive analysis of 2D particle images aimed at increasing the SNR of the data and evaluation of the homogeneity of the sample is performed. (v) An initial low-resolution model of the structure is established using either experimental techniques or computational methods. (vi) The initial structure is refined in order to increase the resolution using an enlarged data set. Only in this step does one attempt to minimize (2.5.7.6) more or less directly. (vii) Visualization and interpretation of the resulting 3D electron-density map is the last step; it often involves docking of X-ray structures of molecules into EM density maps in order to reveal the arrangement of known molecules within the EM envelope (Fig. 2.5.7.1). As within the weak-phase-object approximation of the image formation in EM the relation between densities in collected images and the 3D electron density of the imaged macromolecule is linear [(2.5.7.1)], all data-processing methods employed in the structure determination project should be linear, so the densities in the cryo-EM 3D model can be interpreted in terms of the electron density of the protein.

In the actual single-particle project not all the steps have to be executed in the order outlined above. The technique has proved to be particularly useful in studies of functional complexes of proteins whose base state is known to a certain resolution or even of functional complexes whose atomic (X-ray crystallographic) structure is known. In these cases, steps (iv) and (v) can be omitted and the structure of the functional complex (for examples with ligands bound to it) can be relatively easily determined using the native structure as a starting point for step (vi).

In addition to difficulties with obtaining good cryo-EM data, the technique is computationally intensive. The reason is that in order to obtain a sufficient SNR in the 3D structure, processing of hundreds of thousands of EM projection images of the molecule might be necessary. For each, five orientation parameters have to be determined, and this is in addition to determination of the image-formation parameters required for the optimization of correlation searches. In effect, it is not unusual for single-particle projects to consume weeks of the computer time of multiprocessing clusters. This also explains why the knowledge of the base structure simplifies the work to a large degree: when it is known, initial values of the orientation parameters can be easily established, reducing not only the computational time, but also possibilities of errors in the structure-determination process.

2.5.7.3. Electron microscopy and data digitization

The electron microscope is a phase imaging system; *i.e.*, in order to create contrast in images, they have to be underfocused. Owing to the particular form of the CTF of the microscope [(2.5.7.4)], not only the amplitudes of the image in Fourier space are modified, but information in some ranges of spatial frequencies is set to zero and some phases have reversed sign. Therefore, in order to obtain possibly uniform coverage of Fourier space, the standard practice is to take pictures using different defocus settings and merge them computationally in order to fill gaps in Fourier space. The problem is compounded by the relation between underfocus and the envelope function of the microscope. Far-from-focus images have high contrast, but the envelope function has a relatively steep fall-off limiting the range of useful spatial frequencies. Conversely, close-to-focus images have little contrast, but the envelope function is decreasing, slowly extending useful information to high spatial frequencies. In effect, it is easier to process computationally far-from-focus data and to obtain accurate alignment of particles, but the results have severely limited resolution. Processing of close-to-focus data is challenging and results tend to be less accurate, but there is the potential to obtain high-resolution information.

The experimental techniques of initial structure determination (random conical tilt, tomography) require collection of tilt data. This is facilitated by dedicated microscope stages that can be rotated inside the microscope column yielding additional views of the same field. However, collection of high-quality tilt images is difficult. The quality of tilted images tends to be adversely affected by charging and drift effects. Moreover, as the stage is tilted the effective ice thickness increases (inversely proportionally to the cosine of the tilt angle, so at 60° the factor is two) and the contrast of the images decreases correspondingly. Finally,