

Guest editorial

Image processing and image simulation

Sven Hovmöller

Department of Structural Chemistry, Stockholm University, Stockholm, Sweden

As a method for solving crystal structures, electron microscopy is becoming a serious competitor to X-ray diffraction. The principal advantage of EM is that with electrons we can focus the diffracted beams into an image, something which is not possible with X-rays. Whereas X-ray and electron diffraction patterns only contain amplitude information, electron micrographs also contain the phase information. This is the *raison d'être* for electron microscopy.

While all electron microscopists working on protein crystals know that the phases are present in the images, and routinely extract them from the Fourier transforms of digitised images, it is still a common belief among physicists, and those working with inorganic samples, that the phase information is lost when the image is recorded. As a consequence, very few scientists in HREM apply any image analysis to their images.

How can it be, that what is true in a session of EM in molecular biology is a bad joke in the parallel session of HREM on inorganic materials, in one and the same congress on electron microscopy? One reason is the lack of clear definitions. When physicists talk about phases, they refer to the phases of the diffracted beams *relative to a moving wave*, namely the incident beam. We can call these phases α . When crystallographers, chemists or molecular biologists talk about phases, they refer to the phases β of the Bragg planes (i.e. the planes of electron density in the X-ray case and the planes of potential in the EM case) *relative to a stationary point*, namely the origin in the unit cell. While β is a property of only the specimen, α is a property of the specimen and the measuring device. While the phases α change con-

tinuously as the electron beam works its way down through the crystal, the phases β stay constant. The phases α depend on the structure, and thus on the phases β , but they are lost when the image is recorded.

The electron micrograph of a crystal is a periodic function, and, like all periodic functions, is well suited for Fourier analysis. The calculated Fourier transform of a digitised electron micrograph is a complex function ($A + iB$), and the diffraction points thus have both amplitudes and phases. This is a third set of phases, and we can call them γ .

In order to solve a crystal structure, we need to know the amplitudes and phases β of a sufficient number of diffraction points. Of the phases α , β and γ , only the γ 's are experimentally available. In order to solve the structure, it is necessary to reconstruct β from γ . In an ideal weak phase object and an ideal microscope, γ and β are identical, but in a real specimen γ has been corrupted by multiple scattering, inelastic scattering and electron optical distortions. However, for thin crystals and reasonably well-aligned electron microscopes, the phases γ are very close to the β 's, and it is possible to reconstruct the phases β by crystallographic image processing CIP. Then the crystal structure can be solved.

In addition to the confusion caused by the phase nomenclature, there are other important reasons for the skepticism among materials scientists, against the possibility of obtaining phase information from electron micrographs. The first is the problem of the weak phase object approximation and its breakdown in real specimens. It may be accepted that phases can be restored for

weak phase objects, such as protein crystals, only one unit cell thick. Inorganic crystals, on the other hand, contain mainly strongly scattering metal atoms, and have smaller unit cells, giving rise to multiple diffraction. Such specimens cannot be regarded as true weak phase objects, and therefore (I agree) the phases are no longer exactly the same as they would have been for ideal kinematic conditions.

The second problem is the electron optical distortions (the contrast transfer function, astigmatism and beam tilt) and crystal tilt, which together will distort even an ideal kinematical wave front at the exit surface of the crystal.

The standard procedure for dealing with these problems has been image simulations. Because of the heavy computing involved in image simulations, this method can be mistaken for a quantitative method. Strictly speaking it is not, because the actual comparison between experimental data (the image) and the model is, with a few rare exceptions, done by eye. But this is not the only limitation of image simulations. Perhaps the worst one is that the model structure must be known, in order to simulate an image. Thus, the problem must already be solved in order to solve it! Furthermore, we know from image analysis of HREM images that even the best images are severely distorted, both on amplitudes and phases. These effects come mainly from beam tilt and crystal tilt. In order to keep down the number of image simulations, these two variables are normally set to zero in image simulations. So, the main difficulties in HREM image interpretation are overcome, by defining them as non-existent.

Crystal structure determination consists of two very different parts: solving the structure and refining it. Both logically and experimentally, these two things are completely different, yet they are often confused. In order to solve a structure, all tricks are allowed. One may even guess the structure, but that becomes impossible for structures with more than about five independent atoms.

The refinement of a crystal structure is a different matter. Preliminary atomic positions are known, and from these a calculated Fourier transform gives both amplitudes and phases. In X-ray crystallography, the calculated amplitudes are then

refined against thousands of accurately known amplitudes. If the fit is good, we can be absolutely sure that the structure is correct.

Crystallographic image processing, CIP, is a method for solving crystal structures, but not for refining them. Several complicated and unknown metal oxides have been solved by CIP, but no refinements have yet been performed, so strictly speaking we cannot be absolutely sure that the structures are completely correct.

What, then, is image simulation? It is a unique method for tackling defects, interfaces etc., but it is not a method for solving crystal structures, since the structure model is needed as input to the simulation program. Nor is it a refinement, since there is no least-squares refinement of the model against experimental data involved.

Yet I feel HREM is now very close to the level of perfection of X-ray crystallography as a way to determine crystal structures. We only need to take one last step, the refinement. A structural model for unknown structures can be obtained by CIP. This model should then be confirmed, by a quantitative comparison between calculated and experimental data. This comparison is preferentially carried out in reciprocal space. I can see two ways of doing this. One is analogous to X-ray diffraction; the model obtained by CIP (at some 2.5 Å resolution) is refined against electron diffraction amplitudes to higher (1 Å) resolution. For this we need to improve the techniques for obtaining high-quality (read kinematical) electron diffraction amplitudes from inorganic crystals.

Another possibility is to use CIP for solving the structure and estimating all the electron optical conditions, and then use image simulations to calculate images. The experimental and simulated images can then be compared, in reciprocal space. The structural and electron optical parameters should finally be refined against the amplitudes and phases in the experimental image.

When we have reached this goal, which may not be very far away, electron microscopy will have become a method for determining structures, as accurate and general as X-ray diffraction, but with the much wider field of applications that the EM can offer, since there is no need for millimeter-sized single crystals.