

## ELD – a computer program system for extracting intensities from electron diffraction patterns

Xiaodong Zou, Yuri Sukharev and Sven Hovmöller

*Structural Chemistry, Stockholm University, S-10691 Stockholm, Sweden*

Received 27 July 1992; at Editorial Office 11 September 1992

**Dedicated to Professor Elmar Zeitler on the occasion of his 65th birthday**

A computer program system, ELD, for extracting intensities from electron diffraction (ED) patterns has been developed. ELD runs on a personal computer (PC). Electron diffraction patterns are digitized using a CCD camera, and the data is transferred to the PC via a frame grabber. The lattice vectors and the shape and size of the diffraction spots are first determined, and based upon this information the strategy for extracting the electron diffraction intensities is decided by ELD. It is possible to merge several diffraction patterns taken with different exposure times, whereby both very strong and very weak reflections can be measured. Quantified electron diffraction data can be useful both for chemical applications, such as refining crystal structures, previously solved by crystallographic image processing (CIP), in materials science and for physical applications.

### 1. Introduction

During the last twenty years, high-resolution electron microscopy (HREM) combined with crystallographic image processing (CIP) has been developed and applied for solving structures of proteins [1,2], organic [3] and inorganic crystals [4–6]. CIP is especially useful in the cases when the crystals are too small for X-ray or neutron diffraction. HREM images have the advantage that both phases and amplitudes of structure factors can be obtained, while in X-ray and neutron diffraction the phase information is lost. However, in HREM images the resolution and the quality of the amplitudes are much lower than in X-ray or neutron diffraction.

In order to get higher resolution and more accurate amplitudes, the image data can be supplemented with electron diffraction data. Electron diffraction data goes to higher resolution than images and is not affected by the contrast transfer function. Furthermore, it is easier to align an ED pattern than an HREM image.

Quantitative ED data has been used in several crystal structure determinations the last ten years.

The 3D structure of the membrane protein bacteriorhodopsin was solved to 3.5 Å resolution by combining electron diffraction patterns and HREM images [2]. Phases of 2700 independent Fourier components of bacteriorhodopsin were calculated from the Fourier transforms of the images. The amplitudes were obtained from electron diffraction patterns. An organic structure, perchlorocoronene, was reconstructed by using entropy maximization methods combined with likelihood ranking [3]. Electron diffraction amplitudes were used for extending the resolution from 3.2 Å in the HREM images to 1 Å in the reconstruction. Several crystal structures of lipids and other organic compounds have been studied by Dorset, who also recently reviewed the organic crystal structures determined by electron crystallography [7]. A perovskite-related superstructure,  $\text{Ba}_{28}\text{Fe}_{28}\text{O}_{70}$ , was solved by combining CIP and HREM [6]. A qualitative comparison between the observed electron diffraction pattern and the calculated Fourier transform of two possible structure models helped in identifying which of two possible structure models was correct. Electron diffraction amplitudes to 1 Å resolution were

combined with amplitudes and phases from images in order to find the oxygen positions in  $\text{K}_2\text{O} \cdot 7\text{Nb}_2\text{O}_5$  [8].

Until recently, digitization of electron diffraction patterns was made mainly by microdensitometer. Baldwin and Henderson developed a program system to measure and evaluate electron diffraction patterns to 3.0 Å resolution from two-dimensional crystals of the membrane protein bacteriorhodopsin [9]. The ED patterns were digitized using a flat-bed microdensitometer. Corrections were made for the background, non-linearity of the densitometer, the curvature of the Ewald sphere and other factors. In a later study of the same protein, the possibility of using ED data of heavy-atom derivatives for phase determination of membrane proteins was investigated [10].

Microdensitometers are excellent instruments for digitizing electron diffraction patterns, but they have two disadvantages: they are very expensive, and it is not easy to transfer computer programs developed for one microdensitometer

and the computer that controls it from one laboratory to another, with different equipment. Now that HREM and CIP have become simple tools for solving crystal structures by electron microscopy, we felt it was important to develop a cheap and easy-to-use, yet accurate system for quantitatively evaluating also electron diffraction patterns. The present paper describes such a system, called ELD. (ELD for Electron Diffraction, or Swedish: fire). ELD uses similar hardware (A DAGE MTI CCD72E camera, frame grabber and IBM-compatible PC) as CRISP [11], and it has a similar user-friendly mouse-driven graphics interface. The code for ELD is written in the C language.

## 2. Recording electron diffraction patterns

The metal oxide  $\text{K}_2\text{O} \cdot 7\text{Nb}_2\text{O}_5$  (space group P4bm,  $a = b = 27.5$  Å,  $c = 3.94$  Å) was used to test out optimal experimental conditions and for constructing the algorithm for ELD.

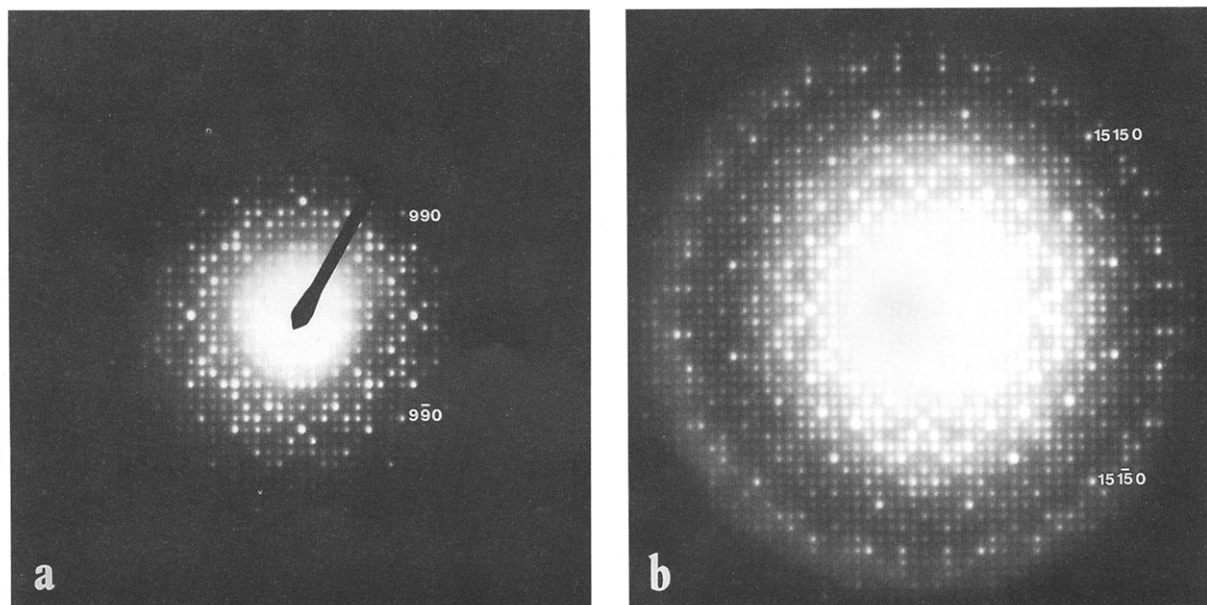


Fig. 1. Electron diffraction pattern of  $\text{K}_2\text{O} \cdot 7\text{Nb}_2\text{O}_5$ . The projection shown is the  $(hk0)$  plane, with  $a = b = 27.5$  Å. Space group P4bm. In order to minimize saturation of the film and CCD camera, the diffraction pattern has deliberately been recorded with slightly blurred reflections. (a) Short exposure time (2.8 s), showing low-resolution reflections. (b) Long exposure time (11 s) showing high-resolution reflections. The indexed reflections (15, 15, 0) are at 1.30 Å resolution.

Samples of the niobium oxide were prepared for electron diffraction in standard ways, i.e. by crushing the specimens in a mortar, and applying the crystal fragments to grids with the help of butanol. Electron diffraction patterns were recorded in a JEOL 2000FX transmission electron microscope, operated at 200 kV. The electron beam was made slightly convergent in order to defocus the diffraction spots.

Diffraction patterns were recorded on Kodak Electron Image film SO-163 and developed in D19 for 10 min.

A camera length of 120 cm was chosen for maximum magnification of the electron diffraction pattern, yet fitting all reflections inside  $1.0 \text{ \AA}$  resolution inside the  $60 \times 90 \text{ mm}$  film.

Several diffraction patterns were collected from each crystal with different exposure times ranging from 0.5 s to over 15 s (fig. 1). In this way the limited dynamical range of the CCD camera could be compensated for. Weak reflections were measured on films with long exposure times, and strong reflections on films with short exposure times.

Since the photographic film has a DQE (detective quantum efficiency) of nearly 1, the optical density of the film is directly proportional to the electron dose, until the saturation is reached. The film saturates at much higher optical density (O.D.) values than the useful range of the CCD camera and thus the O.D. values on the film are considered to be proportional to the number of electrons in each diffraction spot.

Care was taken to align the crystals well and to collect the electron diffraction patterns from thin parts of the crystals, thereby minimizing the effects of multiple scattering. The presence of multiple scattering was monitored as the amount of intensity present in the symmetry-forbidden odd axial reflections ( $h00$ ) and ( $0k0$ ), along the crystallographic screw axes. When only a small area at the edge of a thin crystal was used, the forbidden reflections were very weak.

### 3. Digitizing ED patterns

Digitizing electron diffraction patterns is far more demanding than digitizing electron mi-

croscopy images of crystals for CIP. The reason for this is that each reflection in an electron diffraction pattern is concentrated in a very small area, typically less than 0.5 mm in diameter on the negative. In images, the reflections are the Fourier components in the calculated Fourier transform, with each reflection being present as a cosine wave over the whole image. Thus in images all the pixels contribute to each reflection, while in an ED pattern the information of each reflection is concentrated into very few pixels. For CIP of inorganic crystals it is possible to use the original micrographs, positive prints or even good photocopies of the prints, without significant loss of information.

For digitizing electron diffraction patterns we used only the original micrographs (i.e. the negatives), which were put on a light-box. The contrast was reversed in order to display the diffraction peaks as white and with high numbers, while the background was black and had low numbers in the data files.

The electron diffraction patterns were digitized in the same way as electron micrographs, with a CCD camera, frame grabber and the computer program CCDINPUT as described earlier [11]. The dynamical range present in electron diffraction patterns is very large, from a background of optical density (O.D.) = 0.0 to the direct beam with O.D. > 5.0. With the limited range (256 grey levels) and linearity of the CCD camera, it was necessary to take several ED patterns with different exposure times and to correct the input intensity data pixel by pixel using a look-up table. The look-up table was determined individually for each digitized film, using a calibration film strip from Kodak with 20 known equal steps of optical density from O.D. = 0.05 to O.D. = 4.0. The calibration strip was measured together with the electron diffraction pattern (fig. 2). The response of the CCD camera was close to linear in about half its range of 256 grey-tone values, but fell considerably at the higher readings, as seen from the response curves in fig. 3. The response of the CCD camera could be adjusted in several ways; the zero point could be shifted (from O.D. = 0.0 to 0.6 in fig. 3) and the gain could be adjusted independently. Response

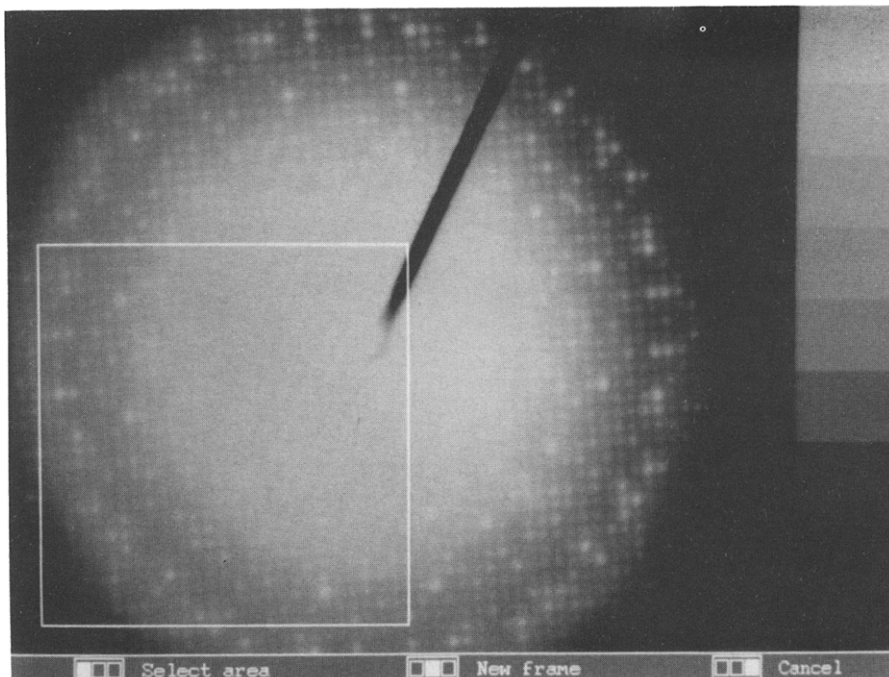


Fig. 2. An electron diffraction pattern has been photographed by the CCD camera, digitized by the frame grabber and is now displayed on the screen of the PC. The whole image is  $640 \times 480$  pixels, and from this one or more sub-areas of  $256 \times 256$  pixels are selected for analysis by ELD. A calibration film strip is recorded simultaneously.

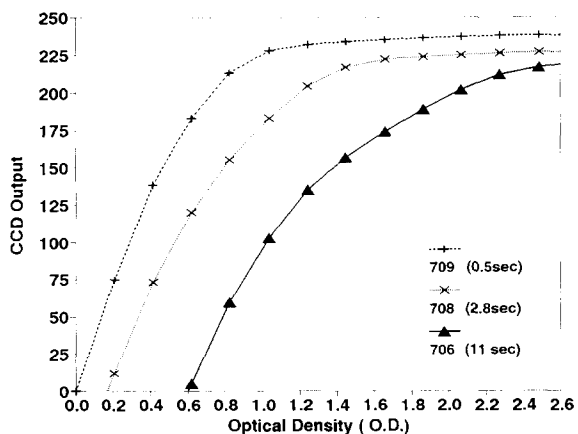


Fig. 3. The response curves of the CCD camera for three different settings. In all cases about half the range of 256 grey levels (8 bits) is a linear function of the optical density. The numbers 706, 708 and 709 refer to different film numbers, and their exposure times (in seconds) are given. The O.D. units are absolute optical density values.

curves for three different settings used for a short, medium and long exposure time, respectively, are shown in fig. 3. The ED pattern recorded in 0.5 s had a background level close to O.D. = 0.0, and so the zero level of the CCD camera was set to O.D. zero, and the gain such that good or acceptable response was obtained up to about O.D. = 0.7. In the ED pattern with longest exposure time (11 s) the background level was as high as 0.6 O.D. units, and that was then chosen as the zero level of the CCD camera. Whatever the zero level was, the useful range of grey-tones covered about one O.D. unit. The precision of the data was always better at lower optical density. In the low O.D. range 0.1 O.D. gave a response of 30–35 digits on the CCD camera, but at one O.D. unit darker, a 0.1 O.D. increase corresponded to only 5–10 digits on the CCD camera.

The look-up table correction is performed by a separate program called IMGTOFFT, which at the same time converts the data format from the image type (IMG) to the format of the Fourier transform (FFT) as used in CRISP. It is imperative to make look-up tables individually for each film, since both the zero level and the gain of the CCD camera is changed from one ED pattern to the next.

Because of the difficulty in handling larger areas than  $256 \times 256$  pixels in a PC, ELD was designed for that size of data files. In the case of  $K_2O \cdot 7Nb_2O_5$  the ED pattern extended to more than 25 diffraction orders in all directions. A complete ED pattern from index  $-25$  to  $+25$  contains 51 reflections in the 256 pixels, giving a maximal possible distance between reflections of only 5 pixels. The accuracy of measuring intensities from such a small spot size is too low. This problem is solved in ELD by splitting up the ED pattern into four quadrants (or even more sub-frames), each with a size of  $256 \times 256$  pixels. The image obtained by the frame grabber (fig. 2) has dimensions  $640 \times 480$  pixels. ED patterns taken with short exposure times show only the low-order diffraction peaks, and all the data in such an

image can be extracted from a single frame, as seen in fig. 4a. ED patterns taken with longer exposure times, extending to  $1.0 \text{ \AA}$  and having over 25 diffraction orders, are extracted as several overlapping frames. One individual frame, covering one quadrant, is shown in fig. 4b. The data from all files constituting a single ED pattern is merged into one single file.

#### 4. Lattice refinement

Before the intensities of the electron diffraction peaks can be extracted, the lattice must be found with sufficient accuracy. In the digitized ED patterns, the  $(0,0)$  reflection (i.e. the direct beam) need not be in the center. This is in contrast to the calculated diffraction patterns obtained as the Fourier transform of electron microscopy images, where the  $(0,0)$  reflection always comes in the exact center of the diffraction pattern. As a consequence, the ED lattice is not determined only by the two lattice vectors, but also the position of the center of the  $(0,0)$  reflection must be refined. The accuracy of the lattice refinement should be such that the predicted

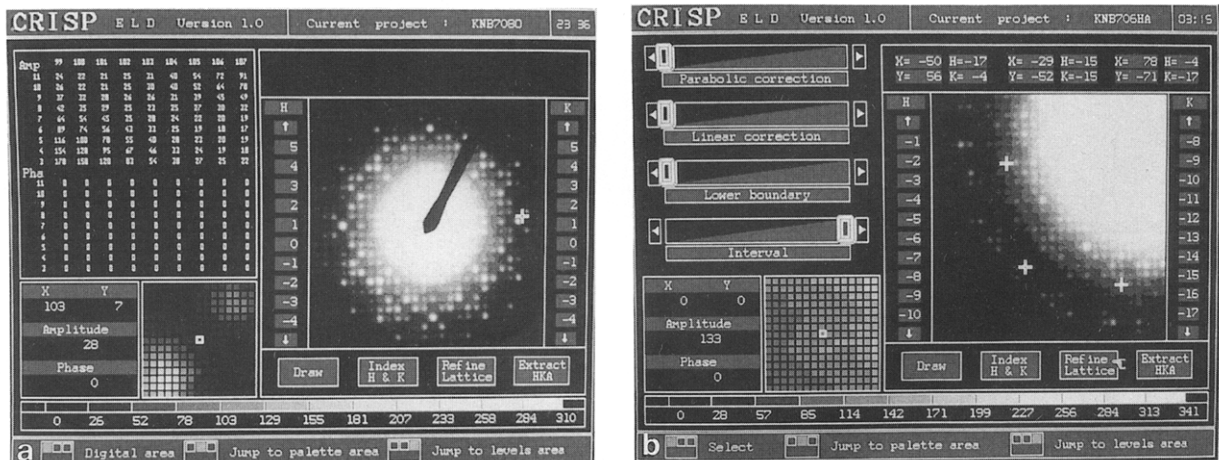


Fig. 4. (a) The ED pattern from fig. 1a displayed by ELD. In the lower left an inset shows a small enlarged area ( $16 \times 16$  pixels) of the image around the cursor position. It can be seen that the reflections are blurred, but not overlapping. In the upper left a  $9 \times 9$  pixel area around the cursor position is shown in digital form. The rapid decline of the very strong  $(10,0)$  reflection about four pixels away from its center is demonstrated by these numbers. (b) One of the four quadrants used for evaluating a film with longer exposure time (fig. 1b), having measurable reflections to very high resolution. Notice that there is no need for having the  $(0,0)$  reflection in the center. Three good reflections are chosen and indexed.

peak positions agree with the experimental positions with an error of no more than about 2% of the distance between the lattice points. In a typical case where the lattice parameters are 10 pixels, this means that the predicted peak position must be correct to within 0.2 pixels. The procedure used for the lattice refinement is as follows.

The user selects three good reflections, points at each of them with the mouse and gives their  $(h,k)$  indices (fig. 4). The criteria for good reflections are: (1) The peak should be strong (with the maximum between 0.3 and 0.9 times the overflow level) and sharp, but not so strong that any pixels reach the overflow level. The overflow level is defined as 90% of the highest pixel value in the frame. (2) The signal-to-noise ratio should be sufficiently high. (3) The averaged background level must be  $> 5$  in order to avoid CCD camera readings under zero which will be set to 0. (4) There should be no very strong neighbours to the reflections.

The computer program first finds more exact  $(x,y)$  coordinates of the three selected peaks, by estimating where the center of gravity of each peak is. The center of a peak in general does not

fall exactly on an integer pixel, but somewhere in between. See for example the reflection (10,0) marked with a cross and enlarged in fig. 5a, with a center at coordinates (98.48, 1.63). Usually the shapes of the reflections follow a Gaussian distribution, and then it is quite straightforward to find the center of gravity.

From each of the three reflections we obtain two equations of the type:

$$H_i Ax + K_i Bx + \text{Orig } x = x_i,$$

$$H_i Ay + K_i By + \text{Orig } y = y_i,$$

where  $i = 1, 2$  or 3,  $(H_i, K_i)$  is the index of the  $i$ th reflection,  $(Ax, Ay)$  is the lattice vector of the (1,0) reflection and  $(Bx, By)$  the (0,1) vector,  $\text{Orig } x$  and  $\text{Orig } y$  are the  $(x,y)$  coordinates of the (0,0) reflection and  $(x_i, y_i)$  are the  $(x,y)$  coordinates of the reflection  $(H_i, K_i)$ .

From the resulting system of equations with 6 unknowns and 6 equations, both the position of the (0,0) reflection and the  $a^*$  and  $b^*$  lattice vectors are estimated. In order to further increase the accuracy of this estimate, ELD searches throughout the entire diffraction pattern for find-

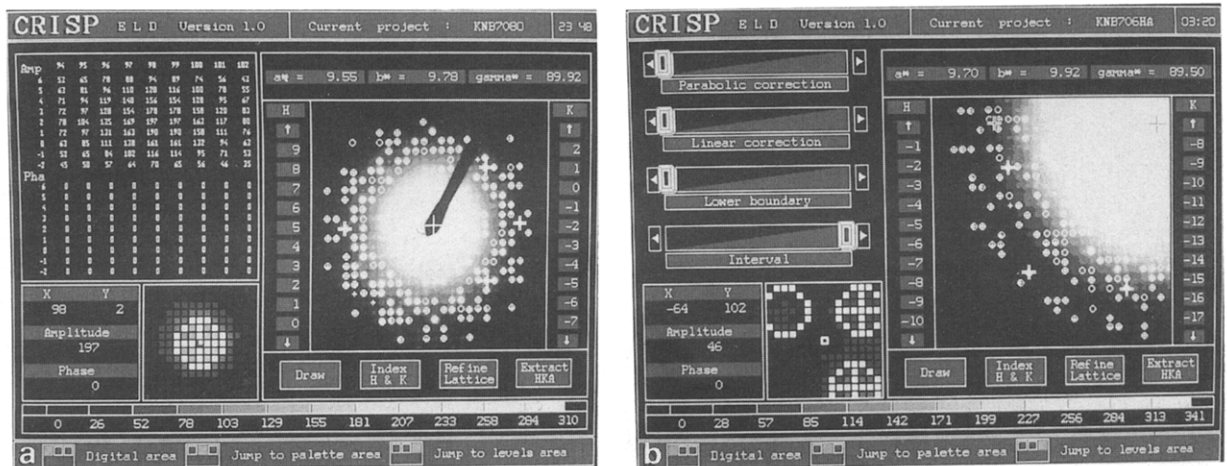


Fig. 5. The reciprocal lattice vectors  $a^*$  and  $b^*$  and the coordinates for the (0,0) reflection have been refined. Only reflections with good quality (neither too strong nor too weak and without too strong neighbours) are used for refining the lattice parameters and the (0,0) reflection. These reflections have been marked with a plus (+) in the first refinement cycle, and with a ring (○) in the second cycle. (a) Film with short exposure time from fig. 4a. Enlarged (10,0). (b) Films with long exposure time from fig. 4b. The different unit cell vectors in (a) and (b) are caused by slightly different magnifications during the digitization. The deviations from perfect four-fold symmetry in the unit cell dimensions ( $b^*$  being 2.3% longer than  $a^*$ ,  $\gamma^*$  not  $90^\circ$ ) are caused by non-uniform magnification in the electron microscope.

ing all reflections ( $H, K$ ) that lie near the lattice  $H(A_x, A_y) + K(B_x, B_y)$  and pass the above mentioned quality criteria, and from the center of gravity of each of these (often some 20 to 200 reflections) makes a least-squares refinement of the 6 parameters. Two cycles of refinement are performed (fig. 5). As a result the coordinates of the (0,0) reflection and the lattice vectors are obtained with an accuracy of around 0.01 pixels, as required.

### 5. Extracting electron diffraction intensities

The procedure for extracting intensities from ED patterns (fig. 6) is quite different from that used for extracting the amplitudes of peaks in the Fourier transform of EM images [11]. In the Fourier transforms of EM images, all peaks in all samples have the shape of a sinc function ( $\sin x/x$ ) (convoluted with the lattice disorder). In electron diffraction the size and shape of the peaks are determined by the camera length and the convolution of the illuminated crystal area and the shape of the beam.

One problem with blurring the diffraction pattern is that adjacent reflections may overlap. Thus there will be an optimum defocus value for each

crystal. This value can be expressed as the ratio between the half-width of the peaks and the shortest reciprocal lattice vector. The half-width of a reflection is defined as the diameter (in pixels) at which the intensity is down to half the value at the center of the peak (after subtracting the background from all values). We have concluded that the optimum defocus is where the half-width of the reflections is near 1/3 of the smallest reciprocal lattice vector.

Since the amount of defocus and magnification will vary from one ED pattern to another, it is necessary to estimate the peak size individually for each film. This is done as follows. Only good reflections are used for this analysis. The criteria for good reflections are the same as those described earlier in the lattice refinement section. For each of the accepted reflections, ELD counts how many pixels have intensities higher than 0.5 times its maximum intensity (after subtracting the background from all values). This number is called the peak size. An overall peak size is calculated by taking the average of all the individual values. The sizes of all individual diffraction peaks are all close to the average size. No tendency of difference in peak size was found between low and high resolution peaks, or between stronger and weaker peaks. In fact the average peak sizes

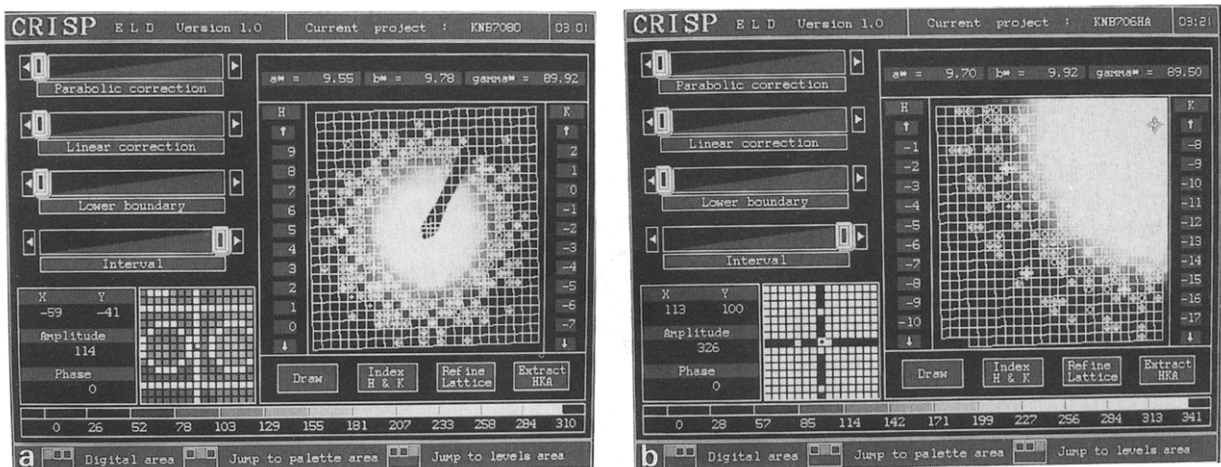


Fig. 6. The ED intensities are extracted by ELD. (a) Intensities are integrated inside a circle at the center of each box, while the background is estimated from the four corners of the box. (b) With the (0,0) reflection in the corner (pixel position (113, 110) marked with a cross) it is possible to extract data out to 1 Å resolution.

obtained were the same (to within 5%) for all the films taken on the same crystal, in spite of the wide range of exposure times from 0.5 to 16 s.

In order to obtain accurate intensities it is desirable to integrate over the whole diffraction peak. This comes into conflict with the risk of running into the tails of adjacent reflections. Also the inclusion of pixels which are very weak compared to the center of the peak is pointless, since such pixels will contribute more noise than signal. We have chosen to integrate the peaks within a circle with a diameter of 1.5 times the half-width as defined above. ELD automatically checks that this value (called the extract radius) is not too big in relation to the lattice dimensions. The extract radius is never allowed to surpass half the smallest reciprocal lattice dimension. The integration of the peaks is done within a circular area around the predicted peak position, using a non-integer extract radius.

Background corrections are done individually for each reflection ( $H, K$ ) by measuring the intensity in the four  $3 \times 3$  boxes situated exactly half-way between it and its four neighbours ( $H - 1, K - 1$ ), ( $H - 1, K + 1$ ), ( $H + 1, K - 1$ ) and ( $H + 1, K + 1$ ). The averaged background value is subtracted from each of the pixels used in integrating the peak. The list of reflections can be edited at any stage; for example it is possible to delete the reflections hidden by the beam stop.

The complete ED data for one structure is collected on several films and each film may have been extracted as several frames. All this data has to be scaled and merged together into one single ED data file. First the different frames from each ED pattern are merged, and then the different films are merged.

The first film scanned was the one with longest exposure time (fig. 1b). On this film the weak high-resolution (1–2 Å resolution) reflections could be seen and measured, but the lower-resolution reflections (> 2 Å resolution) drowned in the high background of inelastically scattered electrons near the direct beam. On the second film, taken with shorter exposure time, many of the stronger high-resolution reflections could still be measured with acceptable accuracy, and these reflections were used to scale the first two films

together. Several reflections were over-saturated on the first film but could be measured on the second film. This procedure was repeated with the following films of decreasing exposure time until all the diffraction spots had been measured.

In the final file with data merged from all films, the individual measurements of symmetry-related reflections are kept. This data can be used for example to analyze the symmetry of the crystal, and to evaluate the quality of the intensity data. It can also be used for other purposes, for example for refining the crystal orientation, but this option is not yet implemented into ELD.

## 6. Evaluation of the quality of the intensity data

The usefulness of the quantitative ED data depends on the number of reflections and their quality. The number of reflections depends on the resolution of the diffraction pattern. If a cut-off is made where the average reflection intensity is equal to the background, then the resolution is typically around 1.0 Å for inorganic compounds. An ED pattern with 1.0 Å resolution contains 6.25 times more reflections than an HREM image with 2.5 Å resolution and in addition to this the ED data might be better, as judged from several criteria.

The total data available is the full 3D data set. A single ED pattern taken along the shortest unit cell axis contains a substantial part of the full 3D data set, but far from all the data. An ED pattern taken along a 4 Å axis contains only about 1/3 of all the unique 3D reflections.

The question of the quality of the measured diffraction points can be divided into two parts. First, how well do the recorded ED patterns represent what we want to measure? A recorded ED pattern is a true representation of the electrons scattered from the sample with its structure, thickness, degree of disorder etc. If the aim is to make a traditional single-crystal structure determination, we want to obtain structure factor amplitudes. Then the sample should be as thin as possible in order to avoid multiple diffraction. It should also be free from defects and perfectly aligned. Some minor misalignment can be toler-



ated since it will be compensated for to some extent when the symmetry-related reflections are averaged. If the aim of the investigation is to evaluate multiple scattering in a quantitative way, the ED patterns from several areas of different but known thicknesses should be analyzed.

Secondly we have the question about how well the recorded ED intensities are being measured. The first criterion for this, the reproducibility, was tested in the following way. One ED pattern was digitized twice, using different magnifications, different settings of the CCD camera and different orientations of the negative in relation to the CCD camera. For the low-resolution part of  $K_2O \cdot 7Nb_2O_5$  as seen in fig. 6a, a total of 426 reflections were measured on both frames. After scaling the two data sets together they had an  $R$ -value of 4.2%. The  $R$ -value is calculated as

$$\sum (\text{Int}(a) - \text{Int}(b)) / \sum (\text{Int}(a) + \text{Int}(b)),$$

where  $\text{Int}(a)$  and  $\text{Int}(b)$  denote the intensities for a reflection measured on both frames  $a$  and  $b$ . This shows that the reproducibility of extracting intensities by ELD is very high. Notice that the  $R$ -values are calculated for intensities rather than for amplitudes ( $F$ -values) as is normally done in X-ray crystallography. If the  $R$ -values were based on amplitudes they would have been about half the values mentioned above.

Intensities of the reflections which were measurable on two consecutive films with different exposure times were compared after scaling the two films together. The  $R_{\text{merge}}$  obtained depended on several factors. The stronger reflections always had lower  $R_{\text{merge}}$  values than the weaker ones. There was no clear tendency when the reflections were sorted according to resolution – sometimes the low-resolution reflections merged better, but in other cases the high-resolution reflections had the lowest  $R_{\text{merge}}$  value. For two films recorded with 11 and 16 s exposure times, an  $R_{\text{merge}}$  of 9.0% was obtained for all 742 common reflections. When the data set from a film exposed only 2.8 s was merged with the two films above, 320 common reflections had an  $R_{\text{merge}}$  of 19.9%. Finally a film exposed 0.5 s was merged with the three films mentioned, and 289

common reflections had an  $R_{\text{merge}}$  of only 7.9%. All the numbers above refer to merging intensity data from the same crystal in the same orientation, before averaging symmetry-related reflections.

Another measure of the quality is the  $R_{\text{sym}}$ , defined as

$$\sum (|\langle \text{Int}(h,k) \rangle - \text{Int}(h_i,k_i)|) / \sum \text{Int}(h,k),$$

where  $\langle \text{Int}(h,k) \rangle$  is the average value of all the  $i$  individual symmetry-related reflections, and the summation is done over all reflections. In the  $(hk0)$  projection of a crystal with P4bm symmetry all reflections occur 8 times, except the axial ( $(h00)$  and  $(0k0)$ ), and diagonal reflections ( $(hh0)$  and  $(h\bar{h}0)$ ), which occur 4 times. The  $R_{\text{sym}}$  depends mainly on the accuracy of the crystal alignment, and it is a very sensitive measure of misalignment. For the crystal shown in fig. 1, which is quite well aligned, the  $R_{\text{sym}}$  is still as high as 20.0% for 2172 measured reflections (315 unique reflections) between 9.0 and 1.0 Å resolution. As a result the estimated standard deviation of the intensities of averaged unique reflections is 7.6% ( $20.0\% / \sqrt{8-1}$ ).

When the reflections were sorted into different resolution ranges the following  $R_{\text{sym}}$  values were obtained: 19.1% for 688 low-resolution reflections (9–3.1 Å), 20.8% for 765 medium-resolution reflections (3.1–2.2 Å) and 25.3% for 719 high-resolution reflections (2.2–1.0 Å).

The reflections were also sorted into different intensity ranges. The 506 strongest reflections had an  $R_{\text{sym}}$  of 16.9%, the 389 reflections of intermediate intensity 24.7% and the weakest 1277 reflections had an  $R_{\text{sym}}$  of 31.9%.

Finally we compared the electron diffraction amplitudes, obtained by taking the square-root of the ED intensities of our structure, with the  $F$ -values (i.e. amplitudes) obtained by X-ray diffraction from the isomorphous but not identical compound [12] with Tl atoms instead of K atoms as in our structure. Since the main framework of niobium and oxygen atoms are similar not to say identical in the two structures, we thought such a comparison could be relevant. The  $R$ -value was 43% for 229 unique reflections

(in the resolution change 3.0–1.0 Å) common to both structure determinations. The differences between these two data sets are due to several factors; the two structures are not identical, one data set is from X-ray scattering and the other is from electron scattering, and finally the experimental errors of the amplitudes obtained by X-ray and electron diffraction.

A qualitative way of evaluating the quality of the ED data is to replace the amplitudes obtained from the Fourier transform of an image of the  $K_2O \cdot 7Nb_2O_5$  crystal, with the ED amplitudes, while keeping the phases from the image. How do the resulting density maps compare? As is seen from fig. 7 the density maps are quite similar. If these two maps are superimposed all the 8 unique niobium atoms coincide. The main differences are that the Nb atoms are sharper and that the K atoms, situated in the 6- and 7-fold tunnels, show up more clearly when ED data is used.

A quantitative comparison of amplitudes from ED patterns and from images was done. The *R*-value for 68 unique reflections common to the two data sets in the resolution range 6.5–3.0 Å was 22.1%.

The amplitudes from the image goes to zero at about 2.5 Å resolution in the image, while the ED data are visible all the way out to 1.0 Å. This could be expressed as a temperature factor, which

is then very much higher for the image than for the ED data. Temperature factor analysis has not yet been implemented in ELD.

## 7. Discussion

Structure determination by crystallographic techniques consists of two distinct parts, solving and refining the structure. For solving a structure relatively few reflections are needed (about 5 reflections per non-hydrogen atom), while for the refinement as many as 50 to 100 reflections per atom is desirable. In order to solve a structure the phases need to be known, and this constitutes the well known phase problem in X-ray crystallography, since the phases are lost in diffraction patterns. In X-ray crystallography it is therefore necessary to estimate the phases in other ways, for example by Patterson techniques or by the so-called direct methods [13]. In electron microscopy the phases are experimentally available, since the diffracted beams are focused into an image. In a previous paper [11] we described the program system CRISP which can be used for determining the phases from HREM images, and solving crystal structures.

The refinement step is very different from solving a structure. A structure is solved when all the atoms, or at least the most strongly scattering

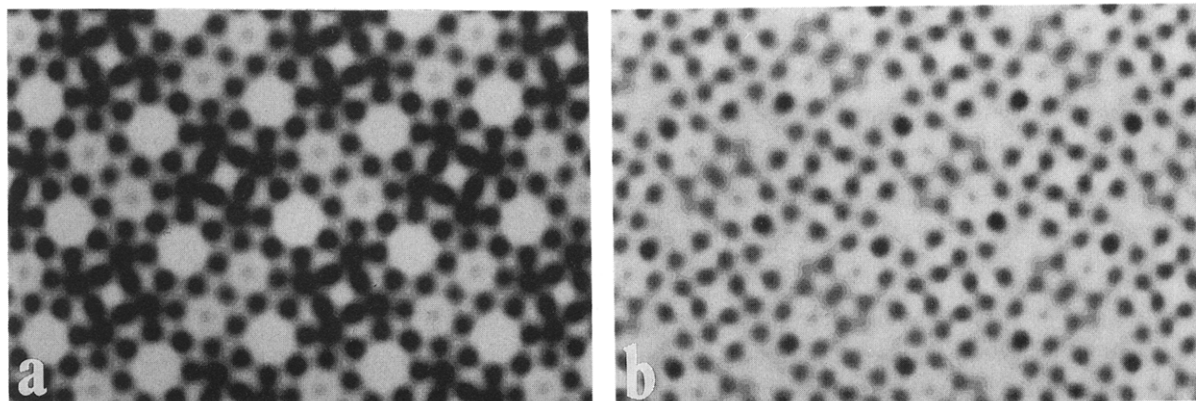


Fig. 7. Density maps of  $K_2O \cdot 7Nb_2O_5$  both produced with the same phases obtained from the Fourier transform of an HREM image of  $K_2O \cdot 7Nb_2O_5$  using CRISP, but with different sets of amplitudes. (a) Amplitudes from an HREM image of  $K_2O \cdot 7Nb_2O_5$ . (b) Amplitudes from electron diffraction obtained by ELD. When the density maps of figures (a) and (b) are superimposed, the atoms fall exactly on top of each other. The atoms are sharper in (b).

atoms, are found with a positional error of less than about 0.1–0.2 Å. The calculated Fourier transform of a model gives both amplitudes and phases. If the structure is essentially correctly solved, the phases of most reflections to 1 Å are known. The structure can then be refined by adjusting the  $(x, y)$  coordinates in order to minimize the difference between observed and calculated  $F$ -values (amplitudes).

The program system ELD is intended for providing experimental data for the second step of a crystal structure determination, the refinement. We have shown here that it is possible to extract intensities from electron diffraction patterns of a thin 3D crystal of a metal oxide with an accuracy comparable to that of X-ray diffraction patterns recorded on film, although not yet comparable to the accuracy of X-ray diffractometer data, which can be collected with extreme accuracy with  $R_{\text{sym}}$  below 1%.

The problems involved in obtaining good ED data are quite different for proteins, organic and inorganic crystals. For proteins [9,10] the main problem is radiation sensitivity, with the additional difficulties caused by bent crystals and the weak scattering due to light atoms and very large unit cells (> 50 Å). Inorganic crystals can withstand electron radiation millions of times better than proteins, but there is the additional problem of multiple scattering. Inorganic crystals diffract to higher resolution than proteins, for several reasons; a 3D crystal is more rigid than a 2D crystal, the smaller unit cell dimensions and higher atomic numbers both contribute to stronger diffraction, and finally the low radiation sensitivity makes it possible to expose the inorganic crystals for longer times, thereby bringing up the weak high-resolution reflections. The total number of reflections to 1.0 Å in the  $(hk0)$  plane of  $\text{K}_2\text{O} \cdot 7\text{Nb}_2\text{O}_5$  is comparable to that of the protein bacteriorhodopsin. Data to 1 Å resolution of  $\text{K}_2\text{O} \cdot 7\text{Nb}_2\text{O}_5$  with a unit cell dimension of 27.5 Å extends to 27 diffraction orders, while the data to 2.7 Å resolution of bacteriorhodopsin [10] with a unit cell of 62.5 Å in the  $p3$  form goes to 23 diffraction orders.

When ED data is to be used for refinement of structures, it should be noted that the relative

scattering factors for X-rays and electrons are different, and that in general the difference in scattering power between atoms with low and high atomic numbers is smaller in electron diffraction than in X-ray diffraction.

The curvature of the Ewald sphere is one of the factors limiting the data that can be collected in a single ED pattern. With 200 keV electrons the Ewald sphere is flat enough to allow data to 1.0 Å to be collected. In the present version of ELD no correction for the curvature of the Ewald sphere is done.

Cooled slow-scan CCD cameras have a much better linearity, with some 4000 grey levels, but they are also at least ten times more expensive than the standard room-temperature CCD camera we used here.

## 8. Conclusions

Electron crystallography may in the future become a very important tool for solving crystal structures from extremely small crystals. In order to bring electron crystallography up to the level of X-ray crystallography, it is necessary to have methods both for solving unknown complicated structures and for refining and confirming the structure models obtained from HREM. This can be done by quantitatively comparing the experimental electron diffraction intensities with those calculated from the model. The computer program system ELD, described here, is designed to measure high-resolution ED data with high accuracy, using relatively cheap and easily accessible instrumentation. The ELD software is available from the authors.

## Acknowledgments

Li Fanghua is thanked for providing the sample of  $\text{K}_2\text{O} \cdot 7\text{Nb}_2\text{O}_5$ . Elmar Zeitler is thanked for his interest in and encouragement of making electron microscopy and electron diffraction more quantitative. Financial support has been obtained from the Swedish Science Research Council (NFR).

**References**

- [1] R. Henderson and P.N.T. Unwin, *Nature (London)* 257 (1975) 28.
- [2] R. Henderson, J.M. Baldwin, K.H. Downing, J. Lepault and F. Zemlin, *J. Mol. Biol.* 213 (1990) 899.
- [3] W. Dong, T. Baird, J.R. Fryer, C.J. Gilmore, D.D. McNicol, G. Bricogne, D.J. Smith, M.A. O'Keefe and S. Hovmöller, *Nature* 355 (1992) 605.
- [4] S. Hovmöller, A. Sjögren, G. Farrants, M. Sundberg and B.-O. Marinder, *Nature* 311 (1984) 238.
- [5] K.H. Downing, H. Meisheng, H.R. Wenk and M.A. O'Keefe, *Nature* 348 (1990) 525.
- [6] X.D. Zou, S. Hovmöller, M. Parras, J.M. González-Calbet, M. Vallet-Regí and J.-C. Grenier, *Acta Cryst. A* (1992) in press.
- [7] D.L. Dorset, *Ultramicroscopy* 38 (1991) 23.
- [8] J.J. Hu, H.F. Fan and F.H. Li, *Ultramicroscopy* 41 (1992) 387.
- [9] J. Baldwin and R. Henderson, *Ultramicroscopy* 14 (1984) 319.
- [10] T.A. Ceska and R. Henderson, *J. Mol. Biol.* 213 (1990) 539.
- [11] S. Hovmöller, *Ultramicroscopy* 41 (1992) 121.
- [12] V. Bhide and M. Gasperin, *Acta Cryst. B*35 (1979) 1318.
- [13] H. Hauptman and J. Karle, *Am. Cryst. Ass. Monograph* #3 (1953).