CRISP: crystallographic image processing on a personal computer

Sven Hovmöller

Structural Chemistry, Stockholm University, S-106 91 Stockholm, Sweden

A computer program system, CRISP, for crystallographic image processing (CIP), has been developed. CRISP runs on a standard personal computer (PC), and is considerably faster than previous systems for CIP. High-resolution electron microscopy (HREM) images are digitized by a CCD camera and directly transferred to the PC via a frame grabber. CRISP has been designed with strong emphasis on user friendliness. Thus, installation takes only a few minutes, a full processing of an HREM image takes less than 20 minutes, all operations are carried out with a mouse, and all the crystallography needed for 2D analysis of images is programmed into CRISP.

1. Introduction

Modern electron microscopes for high-resolution work have a resolution in the range of 0.2 nm, which is near the interatomic distances in molecules and crystals. Yet, HREM images are generally not directly interpretable in terms of atomic arrangement. There are several reasons for this; crystal misalignment and multiple scattering of electrons within the sample and optical distortions caused by imperfect electron lenses are especially serious for inorganic materials. For proteins and other organic specimens radiation sensitivity is the main problem.

In molecular biology, image processing is a standard technique for structure determination of proteins by electron microscopy. After the pioneering work of Henderson and Unwin, who in 1975 determined the 3D structure of the crystalline membrane protein bacteriorhodopsin to 7 Å resolution [1], many molecular biology laboratories have built up their own systems for image processing of electron micrographs. The determination of protein structures to atomic resolution is hampered by several factors, among which low contrast, radiation sensitivity and imperfectness of crystals are perhaps the most severe. After a hard struggle, the main technical problems have now been overcome [2,3] and (near) atomic resolution has been reached for protein crystals studied by EM and CIP [3,4].

Inorganic materials are much easier to handle than proteins. The radiation sensitivity is typically at least a million times less and the contrast several times higher. The smaller unit cell dimensions in inorganic crystals can be both an advantage and a disadvantage. Often there is one unit cell dimension of about 0.4 nm, and images taken along that axis will be free of overlaps of atoms. It is often possible to solve such an inorganic structure from that single projection. With a smaller unit cell, there are fewer diffraction peaks, so each diffracted beam is much stronger than for proteins. On the other hand, the increased multiple diffraction in materials with small unit cells is an important disadvantage.

All taken together, electron micrographs of both inorganic and organic crystals can be much improved by crystallographic image processing. CIP is as useful for inorganic crystals as it is for proteins. Yet, CIP has been used for inorganic applications only in very few cases [5–8]. One of the main reasons for this has been that no simple computer program system for CIP has been easily available. Before starting to analyse images it was necessary to invest in expensive and heavy equip-

ment, such as a microdensitometer and, for example, a VAX computer, and then to spend a year or more setting up the programs in the special environment in each new laboratory. With CRISP, all these difficulties are eliminated. Electron micrographs can be scanned with a small, fast and relatively cheap CCD camera. The whole system CRISP, including both software and hardware, is installed on any 286, 386 or 486 IBM-compatible personal computer in a few minutes. The program is fast and easy to learn and use, also for persons with no previous experience of computers.

The computer code for CRISP has been written mainly in C, except for a few subroutines in Assembler.

Fig. 1. An HREM image of $Cs_xNb_{54}(OF)_{146}$ ($x \leq 8$) [7] has been digitized by the CCD camera. A subarea of 256×256 pixels is selected for further processing.

2. Image input

CRISP handles images, digitized in a size of 256×256 pixels. The digitized images can come from a CCD camera or a microdensitometer. The system includes routines for handling a DAGE-MTI CCD72 camera. An image is captured on a Scorpion 8G frame grabber as 640×480 picture elements (pixels), using the program CCDINPUT. It is possible to correct the magnification independently along the x and y axes of the CCD camera in a range from 0.5 to 2.0 in order to avoid distorted images caused by different magnifications in the x and y directions.

We have used a Nikon AF Micro Nikkor 60 mm, 1:2.8 camera objective lens in front of the CCD camera. With this it is possible to come very close to the image for large magnifications, but also to go to infinite distance for smaller magnification of the image.

Both micrographs and paper prints can be digitized by the CCD camera. This is in contrast to the microdensitometers, which can digitize only transparent film.

From the original image of size 640×480 pixels, a subarea of 256×256 pixels is selected for image processing (Fig. 1). After selecting the area, a histogram of density values within it is shown. This is useful as a check that no pixels fall outside the dynamic range of the CCD camera (under
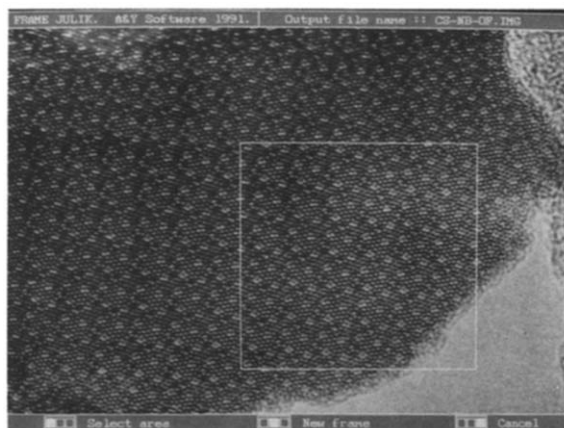
zero or over 256 in this case). A name for the image file is given, and that marks the end of the program CCDINPUT.

3. User interface

All actions except the image input are performed inside the CRISP main program. With very few exceptions, all commands are given by a 3-button mouse. The different options available to the user at each stage are always shown on the display. Therefore it is not necessary to have the manual at hand when running CRISP. A general feature is that the user should not need to burden his memory with names of commands or data files. Within one project (i.e. one image), CRISP assumes that when an action is taken, for example to calculate the Fourier transform, it is on the present image of investigation. It is always possible to go back one or more steps in the procedure, in order to change some parameter. In these cases, the computer remembers the previous parameters used in this project.

When entering into CRISP, a main menu is shown (fig. 2). First a project is selected. This may be a newly digitized image, or an old image, where several steps have been performed already,
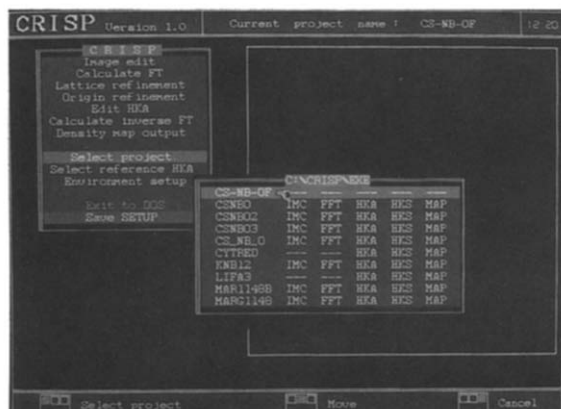
Fig. 2. The main menu of CRISP, showing the different steps of the crystallographic image processing, from Image edit to Density map output, and the selection of an image file (Cs–Nb–OF) from the data base.

Numerical data can be transferred into and out of CRISP at two stages; images and lists of (h, k), amplitude and phase.

4. Editing images

The image, in a size of 256×256 pixels, is displayed with the option Image edit. The images are displayed (fig. 3) with 12 different colours, from the palette of 16 million available with VGA graphics. The user can change individual colours and also define and save up to eight tailor-made colour tables.

Images can be trimmed in three different ways. A circular or rectangular area can be cut out, or an arbitrarily shaped polygon with up to 256 corners may be specified. Statistical information about the image is given in the form of maximum, minimum and mean density values and standard deviations. There is also a histogram of density

and where the user wants to continue the processing or restart from any point with new parameters.
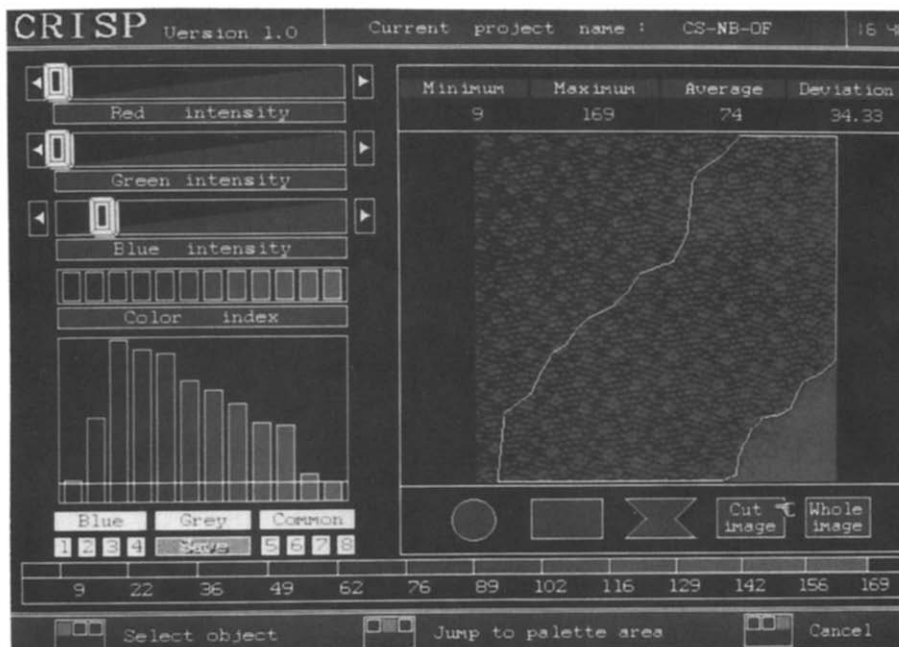

Fig. 3. The digitized image displayed and about to be cut inside CRISP. Statistical information of the image is shown numerically above the image, and in the form of a histogram to the left. The potentiometers red, green and blue can be pulled individually with the mouse, thereby creating any colours desired. Up to eight tailor-made colour tables can be saved, supplementing the three standard tables blue, grey and common.

values. These data are automatically updated to relate to the presently trimmed image. The positional accuracy of the trimming procedure is a single pixel unit, whereby it is possible to separate out areas down to the level of individual molecules or atoms.

The cut option is especially important when the scanned image is not uniform. The two parts of a twin can be cut out and processed individually, for later comparisons. Amorphous regions and crystalline regions on the same image can be processed individually. For wedge-shaped crystals, areas of different thickness can be processed separately, for example in order to follow quantitatively the effects of multiple scattering as a function of crystal thickness.

After an area has been selected, it is automatically stored when the user exits from the Image edit step. The cut image still has the size of 256 by 256 pixels, but the density value of all pixels

outside the selected area has been set to the average density value inside.

5. Fourier transformation

CRISP is intended for use on periodically ordered images, especially crystals. For these applications, the Fourier transformation is a central procedure. A fast Fourier transform routine calculates a 256 by 256 transform of the (cut) image with the same size. The Fourier transform is a complex function on the form $A + iB$. The amplitude part of the transform (i.e. the power spectrum) is calculated as $\sqrt{A^2 + B^2}$ and displayed on the screen (fig. 4). It takes about two minutes to calculate the Fourier transform on a 286 PC equipped with a 287 mathematics processor, and somewhat less time on a 386/387 machine. With-
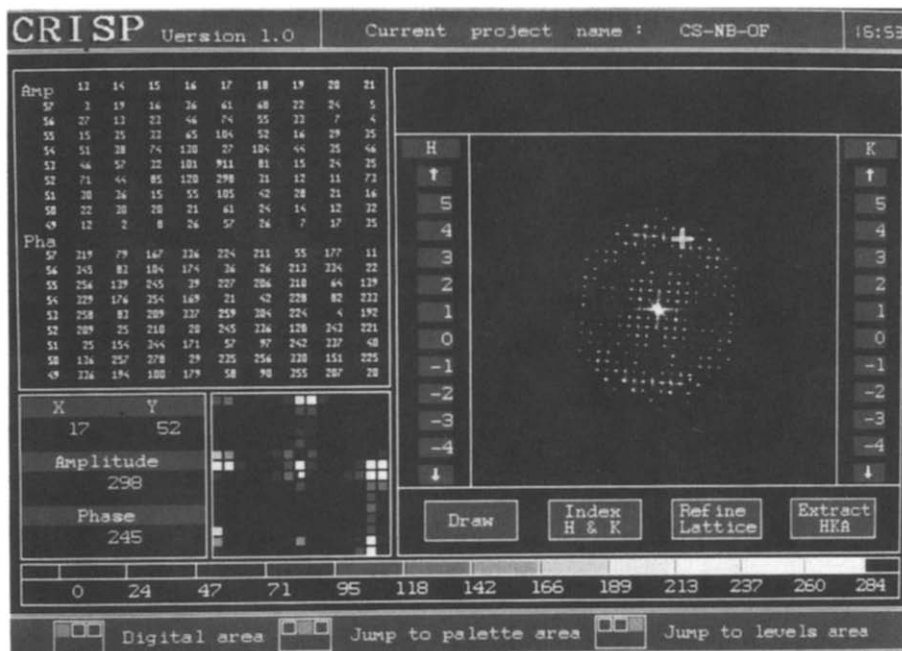

Fig. 4. The diffraction pattern displayed is the amplitude part of the calculated Fourier transform of the cut image in fig. 3. The diffraction pattern seems to show 4 mm symmetry, but the symmetry is not perfect; an effect of crystal tilt. As the cursor is moved over the diffraction pattern, an enlarged image is shown in the magnifying glass at the bottom. The x and y coordinates and amplitude and phase at the present cursor position are shown in the bottom right. It is possible to see amplitudes and phases in a larger (9×9 pixels) area (top left) by selecting Digital area.

out the mathematics processor, the calculation takes ten times longer.

The Fourier transform can be inspected both qualitatively and quantitatively in several ways. With the potentiometer Interval the diffraction pattern can be made darker or brighter. Already by visual inspection of the diffraction pattern it is possible to get some idea of the symmetry of the crystal. It is critical for the success of CIP to know the correct projection symmetry of the object. In this case (figs. 4 and 5) the crystal seems to show both 4-fold and mirror (mm) symmetry, although the symmetry is not perfect – note how all reflections at the top and bottom are stronger than their symmetry-related ones at the left and right sides. In later steps of CRISP the symmetry properties of both amplitudes and phases are evaluated quantitatively.

An enlarged area around the present cursor position is shown in the magnifying glass at the bottom of the screen. The numerical phase and amplitude values and x, y coordinates at the pixel position under the cursor are continuously shown on the screen. There is also an option for looking at amplitudes and phases in a larger (9 by 9 pixels) area (fig. 4 top left) by the option Digital area. In this way inter alia the shape of individual diffraction points and the background noise level can be inspected.

The purpose of CRISP is to perform crystallographic image processing, CIP. For this it is essential to first find and refine the lattice of diffraction points, and then to extract the amplitudes and phases of every diffraction point that falls on this lattice. The lattice is defined by the user, by pointing at two diffraction maxima and specifying their indices (fig. 5). All other strong diffraction points are found automatically, and after two cycles of least-squares refinement the exact lattice is established. This is a critical step, and the algorithm for finding the lattice has been written to eliminate the possibility of erroneously
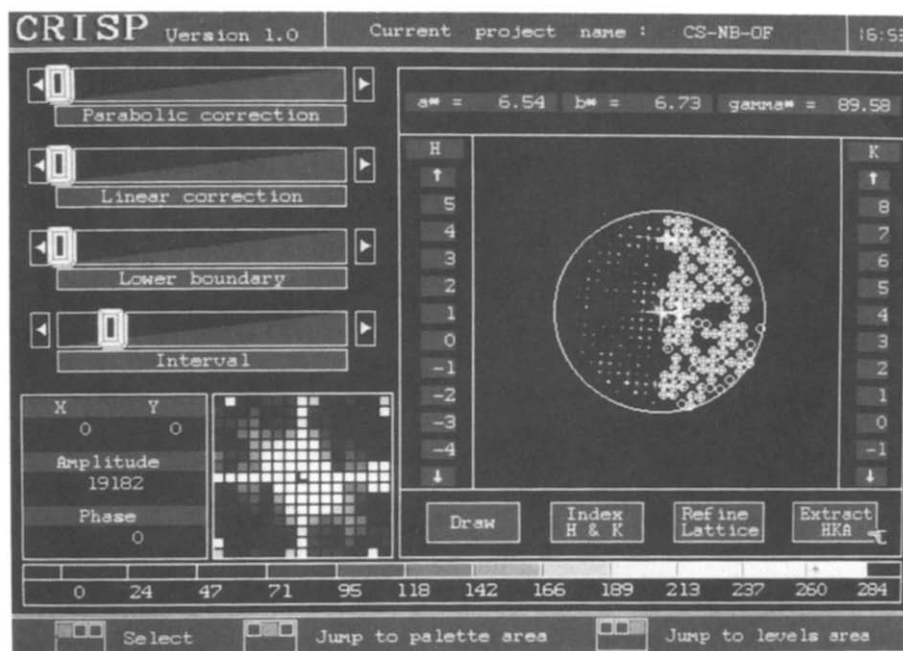

Fig. 5. The diffraction pattern has been indexed, by pointing at two diffraction points and specifying their indices. In this case (2,0) and (0,8) have been selected and marked by white crosses. With the option refine lattice the program finds all clear diffraction points in concordance with the first two reflections, and calculates a new lattice with a least-squares refinement procedure. After another cycle of refinement, the lattice is well determined and the lengths of the reciprocal lattice vectors (in pixel units) and angle are displayed. After this the amplitudes and phases of all diffraction points are extracted by Extract HKA.

including strong peaks coming from the background noise and not the crystal. The lattice must be determined with such high accuracy that even the highest diffraction orders that are going to be extracted will be found with a positional error of less than half a pixel. This positional accuracy is needed for extracting correct phases, because the phase value changes by 180° one pixel away from the center of the peak.

Once the lattice is known exactly, amplitudes and phases are extracted in the same way as in our previous image processing system [5], which in turn was based on the system of programs developed at the MRC Laboratory of Molecular Biology in Cambridge [9]. Amplitudes are extracted by integrating the 3 by 3 pixels closest to the predicted lattice peak position (which in general of course does not fall exactly on a pixel) and then subtracting the average background at the edge of a 9 × 9 pixel box around each diffraction peak. The phases of the reflections are taken as the phase of the pixel closest to the predicted center of the peak. This accuracy is sufficient, since the phase is constant within one pixel from the center of the peak. However, just outside this limit, the phase is reversed i.e. changed by 180°.

In order not to have peaks coming into the background of their neighbours, the magnification of the image should be chosen such that both lattice vectors are at least 6 pixels long. This is easily accomplished by selecting an appropriate distance between the CCD camera and the image during the input. The length of the lattice vectors (in pixel units) is the same as the number of unit cells that fit into the scanned image area. Thus, if 6 by 20 unit cells of a certain crystal fill the 256 × 256 pixel image area, then the lattice vectors will be 6 and 20 pixels long, respectively. Since the radius of the diffraction pattern is 128 (half of 256), it will be possible to extract diffraction points up to the 21st and 6th orders, respectively. For all normal purposes we have found the 256 × 256 format to be sufficient, but for extremely good images, where more than 20 diffraction orders are expected, it will be necessary to process images of larger sizes, for example 512 × 512, which is not possible with the present version of CRISP.

6. Symmetry determination

The main power of CIP lies in finding and imposing the correct symmetry of the crystal. The simple filtering procedure called lattice averaging gives a significant improvement only of very noisy images, such as those of organic molecules or proteins. HREM images of inorganic crystals are, however, virtually free of this white random noise, and thus are not significantly improved by averaging over many unit cells. For inorganic crystals the main distortions of the images are caused by systematic errors, such as crystal tilt and electron optical distortions, affecting all unit cells in the same way. In most cases it is possible to reduce or even eliminate these systematic errors, by imposing the correct crystal symmetry on the data.

Symmetry elements cause relations among the amplitudes and phases of symmetry-related reflections. For example the two reflections (h, k) and $(h, -k)$ have equal amplitudes for orthorhombic crystals. In X-ray crystallography the symmetry determination is reliably done only on amplitude relations, but this is not so suitable for electron microscopy. There are two reasons for this; in X-ray diffraction the scattering process is nearly ideally kinematical, and so the diffraction intensities can be very accurately determined, while electron scattering is much stronger and thus the kinematical approximation breaks down in the EM. Even more important is, however, the very strong effect on amplitudes that even a very tiny crystal tilt will have. Typically, the amplitudes of symmetry-related reflections differ by a factor of 2 or more even for very good HREM images of thin inorganic crystals. A measure for the quality of amplitude data much used in X-ray crystallography is called R_{sym}. R_{sym} shows the difference of amplitudes of symmetry-related reflections, and is defined as

$$R_{sym} = \frac{\sum_{hkl} ||F(hkl)| - |F(h'k'l')||}{\sum_{hkl} |F(hkl)|}$$

where (hkl) and $(h'k'l')$ are symmetry-related reflections.

In X-ray crystallography R_{sym} is usually below 5%, whereas for electron microscopy data R_{sym} is typically around 50%.

The distinction between rotation axes and screw axes is done in X-ray crystallography by inspecting the presence or systematical absence of reflections with odd indices along the crystal axes. Because of multiple scattering, systematically forbidden reflections can often be present in the EM data. For the above reasons, it is not very easy to determine the symmetry of crystals only from amplitudes obtained by electron microscopy or electron diffraction.

Fortunately, the phases are generally of a much higher quality than the amplitudes in electron micrographs. This is in contrast to X-ray crystallography and electron diffraction, where the phases are not experimentally available. From the relative phases of the different reflections, it is possible to uniquely determine the symmetry for most good HREM images of inorganic crystals.

While there are 230 different 3D space groups, there are only 17 plane groups. EM pictures are 2D projections of crystals, and so there are only the 17 2D symmetries to distinguish among for HREM images of crystals. For the nomenclature of the symmetries, we have followed the one commonly used in electron crystallographic studies in molecular biology [9]. These were first derived by Holser [10], for the 80 two-sided plane groups or layer groups, i.e. 2D plane groups with the addition of a distinct top and bottom side, or the 3D space groups that are only one unit cell thick. Out of the 80 two-sided plane groups, only 17 are possible for chiral molecules like proteins, since they are the only ones which do not contain inversion centers, mirrors or glide planes. The 17 chiral two-sided plane groups stand in a one-to-one relation to the 17 plane groups [9]. The niobium oxide used as a demonstration of CRISP in this paper has the 3D space group P4bm. The projected structure down its 4-fold axis is plane group p4g, and according to the nomenclature commonly used in electron crystallography it is p42$_1$2.

The 17 symmetries can be distinguished by the different phase relations they give rise to. For example, a crystal projected along a 2-fold axis will give rise to a centro-symmetric image, which is characterized by all reflections being either 0° or 180° (provided the phase origin of the unit cell coincides with a 2-fold axis, and that the noise is zero). As higher and higher symmetries are explored, more and more relations are found. In p4, not only must all reflections be 0° or 180°, but pairs of reflections (h, k) and $(k, -h)$ should also have equal phases, as a consequence of the 4-fold symmetry. In p42$_1$2, in addition to the phase relations for p4, also pairs of reflections (h, k) and $(h, -k)$ are related. They have equal phases if $(h + k)$ is an even number, and phases differing by 180° if $(h + k)$ is odd. p42$_1$2 is the symmetry of the crystal used in this paper.

All the phase relations for each of the possible 17 projection symmetries are programmed into CRISP. If the symmetry of the crystal is not known before, it may be found out by running the origin refinement routine for all the possible symmetries and comparing the resulting phase residuals.

7. Origin refinement

The procedure Origin refinement looks for symmetry elements in the image, by an analysis of the phases of the reflections.

Before running the origin refinement, the reflections that should go into the calculation are selected, using the option Radii (fig. 6). Only the reflections between the inner and outer radii are included in the origin refinement calculation. Most often the inner radius is set to zero, and the outer radius is set to just include the highest strong reflections. The outer limit is usually very easy to choose, because the amplitudes tend to fall off drastically at a certain resolution, which incidentally usually is well inside the Scherzer focus limit.

The origin refinement routine is based on the principle already implemented in the MRC series of programs [9]. Every point within one unit cell is tested for the fulfillment of the phase constraints given by the symmetry. If the symmetry p21 is selected, CRISP tests how close every reflection is to the nearest of 0° and 180°, at every

point in the unit cell. If the data is perfect and noise free, there will be four points in each unit cell where all reflections are either $0°$ or $180°$. These are the centers of symmetry. In most places of the unit cell, however, the phases are randomly distributed between $0°$ and $360°$, with an average deviation of about $45°$ from the nearest of $0°$ and $180°$. For real pictures, there is always some noise and the data is not ideal, but still there is usually four very distinct positions in the unit cell where the phase residual is about $20°$ or better (lower) (fig. 7a): These four points are the positions of the centers of symmetry.

Once a good symmetry center is found the origin of the unit cell is shifted to be on that position, simply by selecting Shift phases (best). The rule-of-thumb is that if a phase residual of less than $20°$ is found, then the image is good enough to be used for further processing. Phase residuals over $30°$ are usually too large to be possible to compensate for by CRISP. The phase residual also depends on the number of reflections included (rising with more reflections) and the symmetry (rising with higher symmetry, since more and more conditions have to be fulfilled).

Usually the reflections are amplitude weighted, but if equal weights are used, then the phase residual will be higher, since the weaker reflections then get the same weight as the stronger reflections, although their phases are less accurately determined.

In order to get a feeling for the result, the output of the origin refinement is in the form of a colour coded map of the phase residual within one unit cell. The best (lowest) phase residuals are red, followed by yellow, green, blue and black. CRISP automatically finds the best position for the phase origin, i.e. the position in the unit cell with lowest phase residual, and the phase residual at that position can be read out on the display.

The symmetry of the niobium oxide projection was investigated by calculating the best phase residual for the different possible symmetries. The phase origin of the crystal was first refined using p21 symmetry (fig. 7a). The four equivalent centers of symmetry (the red areas) gave a phase residual of $16.3°$, for the 76 reflections to 3.8 Å resolution. This shows that the image is centrosymmetric. For p4 the phase residual was $19.2°$.
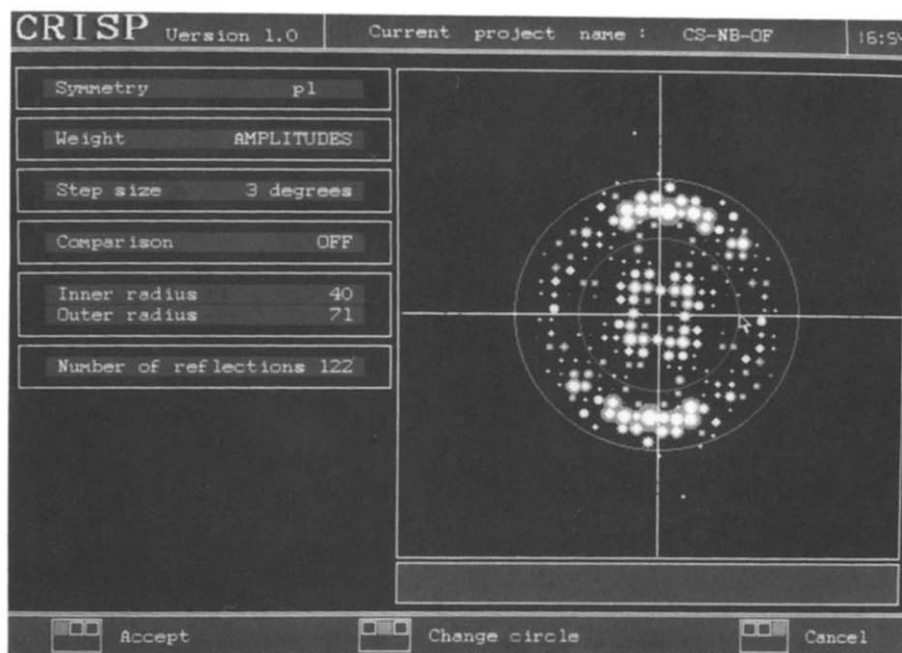

Fig. 6. Reflections between an inner radius (arrow, circle) and an outer radius (circle) are selected for origin refinement.
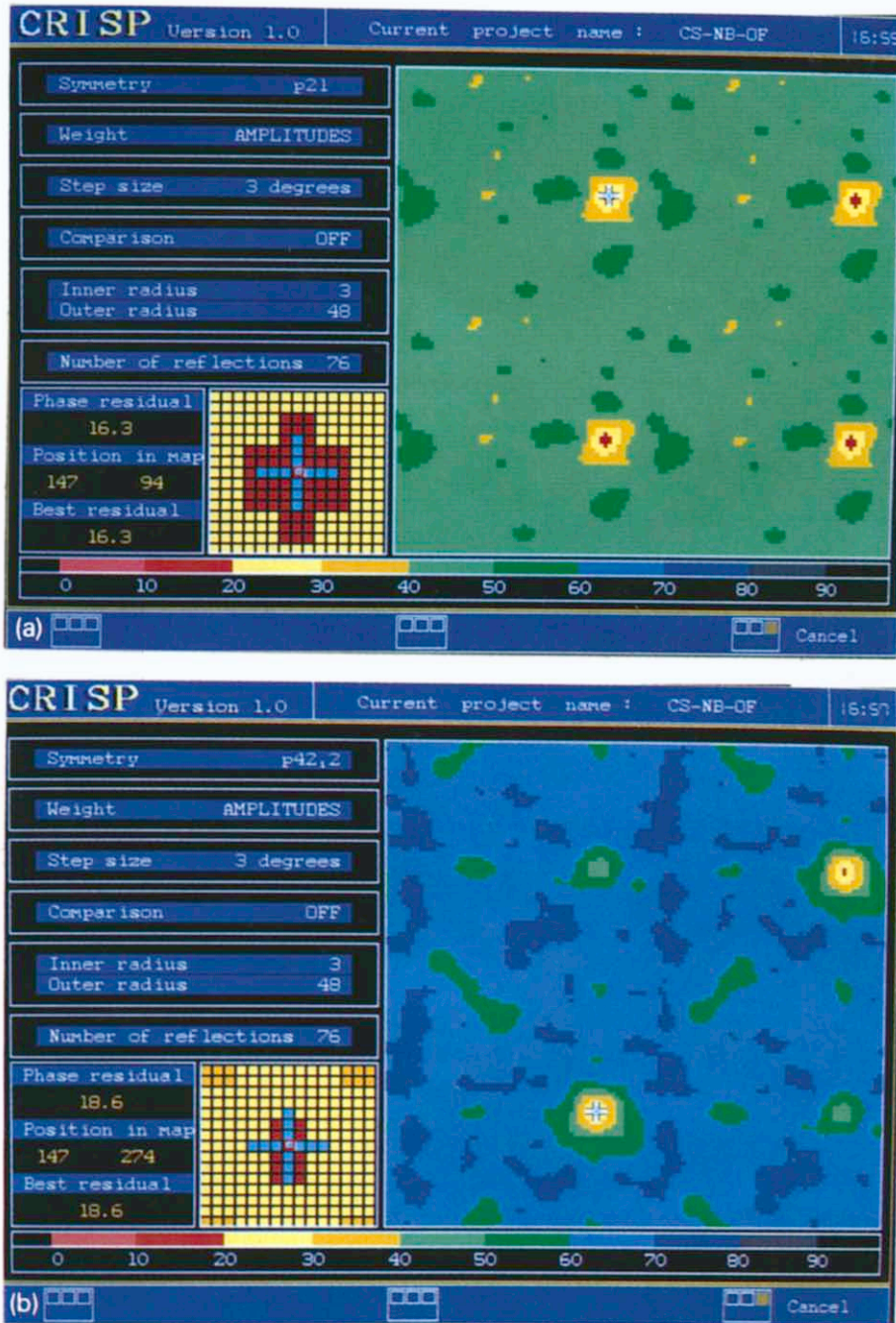
Fig. 7. Phase residual maps for origin refinements in different symmetries. The phase residuals are colour-coded. The best phase residuals (0°–20°) are red. (a) For p21 there are four equivalent origin positions, all with a phase residual of 16.3°; while in (b) p42₁2 there are only two equivalent origins. The best origin positions in p42₁2 have nearly the same phase residuals (18.6°) as those in p21, while other positions in the unit cell have higher phase residuals.

Since the phase residual was almost as good in p4 as in p21, in spite of the much stricter conditions on the phases in p4, it can safely be concluded that the crystal is tetragonal, with at least p4 symmetry. In order to check the possible presence of further symmetry elements, also the other two tetragonal symmetries p422 and $p42_{1}2$ (fig. 7b) were tested. The phase residuals were 39.2° and 18.6°, respectively. This shows clearly that the symmetry is $p42_{1}2$.

Once the symmetry of the crystal has been found, the further processing is done in that symmetry.

It is also possible to compare the phases of two different images with each other, within the origin refinement procedure. This may be useful for areas with different thicknesses within one crystal, or for images of the same area taken with different focus, or even for comparing two different crystals.

8. Imposing crystallographic constraints

Once the symmetry of a crystal has been found, the constraints on amplitudes and phases can be imposed by CRISP within the step Edit HKA. It may be repeated that it is crucial that the symmetry is correctly determined. Fortunately, the result of choosing an incorrect symmetry will almost inevitably be a poorer density map. If a too low symmetry is applied (for example p2 in stead of p4), then the map will be essentially correct, but more noisy than necessary. If a too high symmetry is applied, unrelated features are averaged, and the result will be a more blurred map, which is harder or impossible to interpret, compared to the correct density map.

When the symmetry has been correctly determined, it is also imperative to find the position of the symmetry elements within the unit cell. The symmetry-averaging of the phases can only be performed if the phase origin is correctly positioned with respect to the symmetry elements. In CRISP the symmetry elements are on the same functional form (i.e. have the same origin) as those in the corresponding plane groups, as listed

in the International Tables for X-ray Crystallography [11].

In the lowest symmetry, p1, there are no restrictions on phases, and no symmetry-related reflections (except for the trivial relations between Friedel pairs; (h, k) and $(-h, -k)$ have the same amplitudes, and phases α and $-\alpha$, respectively). Therefore, the amplitudes and phases resulting from choosing p1 symmetry are just the raw data, after lattice averaging.

In p21, there are still no symmetry-related reflections other than the Friedel pairs, but the phases are restricted to either 0° or 180°. When p21 is selected, phases between $-90°$ and $+89°$ are set to 0°, and the phases in the range $+90°$ to $+269°$ are set to 180°. The phases that are changed more than 45° are marked as red lines on the screen, and they also get the flag P, indicating a problem with the phase.

In p4 all reflections are centro-symmetric, but further pairs (h, k) and $(k, -h)$ must have equal amplitudes and identical phases. CRISP sorts such pairs of symmetry-related reflections together in the list of reflections, gives both reflections an amplitude which is the average of the two, and gives both reflections the same phase. The phase value given to the two reflections is derived through an algorithm which takes both the phases and amplitudes into account. First the two reflections are summed as complex vector sums $(A_{1} + iB_{1}) + (A_{2} + iB_{2})$, and then the phase given to both will be 0° if $(A_{1} + A_{2}) > 0$, otherwise the phase is set to 180°.

In $p42_{1}2$ (fig. 8) the four symmetry-related reflections (h, k) $(h, -k)$, (k, h) and $(k, -h)$ are grouped together. They will all be given an amplitude which is the average amplitude within the group. The phases they get are calculated using a procedure similar to the one described for p4, but with the extra condition that if $(h + k)$ is odd, then only (h, k) and $(k, -h)$ should have identical phases, ϕ, while $(h, -k)$ and (k, h) both have phases $\phi + 180°$. When $(h + k)$ is even, then all four reflections in a group have identical phases.

All the 17 plane group symmetries have been programmed into CRISP, each with their special

characteristics regarding symmetry-related reflections and phase relations.

Whenever there are problems, these are marked as red lines in the list of reflections. The kind of problem is marked with flags P, A or *. If the phase of a reflection is more than 45° away from that which has been deduced for the group of reflections to which it belongs, it is flagged with a P. If the amplitude of a reflection has been measured to be negative (i.e. the background was higher than the values under the peak), it will be flagged A. Finally, such reflections which should be systematically absent in the symmetry specified, yet have been measured to have a positive amplitude, are indicated by an asterisk.

It is also possible to edit individual reflections by hand. This can be useful, for example, when a group of reflections is poorly phased, or for deleting too uncertain reflections. If the focus is such that some reflections fall outside the first crossover of the contrast transfer function, these reflections will have suffered a phase shift of 180°,

which should be corrected by manually subtracting 180° from their phases. In practice we have found that the automatic phase determination is so powerful that normally we do not feel any need to edit the data by hand. The few high-resolution reflections that have poorly determined phases usually have such weak amplitudes that their contribution is negligible anyway, so the density map does not change significantly by deleting them.

There are also options for changing the cell parameters, if they are more exactly known from other measurements. Typically orthorhombic and tetragonal unit cells will come out with cell angles up to one or two degrees away from 90°, and this should then be set to exactly 90°.

9. The density map

A density map is calculated from the edited reflection data, by an inverse Fourier transforma-

Fig. 8. In Edit HKA the crystallographic constraints are imposed on amplitudes and phases. Out of the 17 possible symmetries, p42₁2 has been selected here. Symmetry-related reflections are grouped together, their amplitudes (AMP) are averaged (sAMP) and the best allowed phase values are imposed (sPHA). Notice the poor quality of amplitudes, but the very good phases, a typical feature in HREM images.
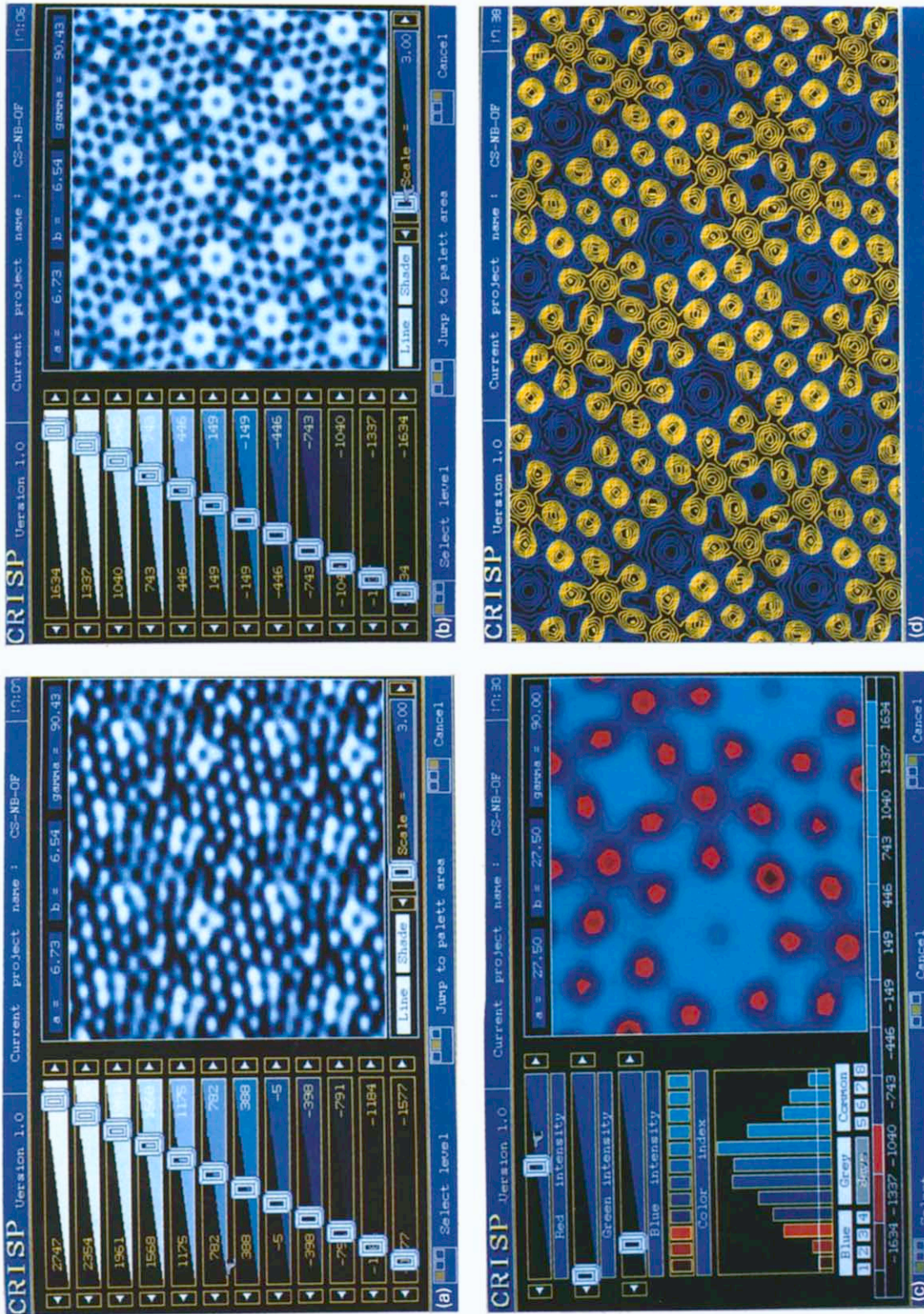
Fig. 9. (a) Density map with p1 symmetry. Imposing p1 symmetry is equivalent to performing only a lattice averaging. In this case, as is typical for HREM images of inorganic crystals, the random noise is very small, so the different unit cells are practically identical, and nearly nothing is gained by only lattice averaging. (b) Density map after imposing the p4₂2 symmetry. The scale is the same as in (a). A very substantial improvement is evident. Now all niobium atoms (black dots) are well resolved, and have a nice round shape. (c) Part of the unit cell after applying p4₂2 symmetry, enlarged. The two highest levels of density have been made red, to emphasize the very uniform size of the different niobium atoms after CIP was applied. Notice the regular 7-fold tunnel with a partially occupied Cs in the center, and the regular pentagon to the right. (d) The density map displayed with contour lines and shown over the full screen.

tion. Strictly speaking it is not a density map, but the projected potential; yet in fact these two things appear very similar. It is advised always first to calculate a map with p1 symmetry (fig. 9a), in order to have the raw image data as a reference. In the map after imposing the crystallographic constraints on the phases and amplitudes (fig. 9b), the structural details are much easier to distinguish. Atoms that merge in the p1 map are well resolved in the higher symmetry. Chemically identical atoms often come out with very different intensities in p1 maps, but are more alike in the maps calculated with the full symmetry.

From the symmetrized density maps, it is possible to localize the atomic coordinates of metal atoms in oxides, with an accuracy of 0.1 Å or better [5–7].

The scale of the maps can be changed over a wide range in CRISP. Some ten-by-ten unit cells are shown on the screen using the lowest magnification. The highest magnification, being 25 times higher, shows just a part of the unit cell on the whole screen.

It is possible to change the colours of individual density ranges (fig. 9c). It is also possible to change the levels of the different contours, although these are usually kept at the default values, where the density is divided into 12 equidistant bins, going from the lowest to the highest density values. An alternative representation, with contour lines, is shown in fig. 9d.

Apart from being displayed on the computer graphics screen, there are options for making hard copies on paper. The density maps can be output by either laser printers or 8-pen plotters. It is also possible to get laser printer outputs of the diffractograms and images, both before and after trimming.

10. Comparison with other image-processing programs

CRISP is built on the philosophy of our previous image-processing system [5], which ran on a VAX 11/750, and took images from a Joyce Loebl microdensitometer MDM6. That system

was written in FORTRAN, and was an implementation and extension of the image-processing system from the MRC Laboratory of Molecular Biology in Cambridge [9]. However, CRISP is completely rewritten from scratch, in the computer language C.

The most dramatic improvement, in terms of time, is the digitization of images. With the microdensitometer it took about an hour to scan an image, while the CCD camera only needs a few seconds. We were afraid that the quality of the CCD camera would be so poor, in comparison to the microdensitometer, that the data quality would be impaired. The dynamic range is only 256 grey-levels on the CCD camera compared to 2000 on the microdensitometer. However, we have not been able to see any deterioration, neither on amplitudes nor on phases, after switching to CCD camera.

The time needed for processing an image is less than 20 min with CRISP (as can be seen from the clock in the upper right-hand corner of figs. 3 to 9), compared to some 5 h with our previous system. The new system is also very much easier to learn to use than the old one. In CRISP the available commands are always clearly indicated on the screen, and executed by a click on the mouse. In our previous system, some procedures were interactively started by writing command lines, while others were run as batch jobs. The origin refinement map used to be printed on paper by a line-printer in an undigestible form. With the colour-coded output in CRISP there is a chance to learn something from the phase origin maps – they show some interesting patterns, related to the structure, which we have not yet even started to exploit.

Another major improvement is the time needed for implementing the image-processing system in a new laboratory. The old system was tailor-made for a special microdensitometer, computer and graphics terminal, and large portions of the program code had to be rewritten before it could be used on another VAX computer. With CRISP it is sufficient to make a copy of the programs from a floppy disk onto the hard disk of a personal computer, and then the whole system is ready to use. To install the CCD camera it is only neces-

sary to replace the VGA card of the personal computer with the frame grabber card. The whole installation of CRISP, including the CCD camera and the frame-grabber card can be done in about ten minutes.

Because the computer keeps track of all the intermediate results, the user need not do that, and can concentrate on the scientific part of the image processing, rather than the tools. The speed and comfortable, not to say enjoyable, use of CRISP encourages the user to experiment and to process larger series of images than were feasible earlier.

The CRISP image-processing system is now commercially available from Parameter, Grevgatan 39, S-114 53 Stockholm. Several other image-processing programs are commercially available and in frequent use for electron microscopy applications. Among these are SEMPER [12] from Synoptics, IBAS and VIDAS from Zeiss-Kontron, IMAGIC [13] from Image science and SPIDER [14]. All programs have some features in common, but they also each have their own specialities. All the systems have options for calculating and displaying Fourier transforms, but the filtering algorithms for extracting the amplitudes and phases from the transform are not as easily available as in CRISP. In some cases, the filtering algorithms are directly erroneous for crystallographic applications, for example when the filtered data is chosen as the highest peaks close to the predicted lattice. As mentioned above, the phases are reversed (shifted 180°) only one pixel away from the actual peak position, and so any phase information not extracted strictly on the predicted lattice points will be incorrect.

The crucial phase origin refinement step is only commercially available in CRISP, although similar procedures are standard in many local versions of image-processing systems, most of which are based on the Cambridge system.

11. Discussion and conclusions

It is our hope that the efficiency of CRISP will help to promote the use of CIP in many new fields of electron microscopy, especially in inorganic chemistry and materials sciences, where it

has not been much used yet. Applications can be found in mineralogy, geology, solid state chemistry, crystallography and many other fields.

It would be interesting to use this type of image analysis as a way to quantitatively study the effects of multiple electron scattering through different materials. The data in the actual experimentally produced electron micrographs are conveniently extracted by CRISP, and could be compared to those derived from theoretical calculations by image simulations [15]. It is already clear from our findings up to now that the two main reasons why HREM images of thin crystals are not directly interpretable are factors which make the amplitudes and phases lose the symmetry that is present in the crystal [7]. These factors are mainly beam tilt and crystal tilt. The serious effects on electron micrographs of beam tilt have also been found from theoretical image calculations [16]. Image simulations are often done only varying defocus and crystal thickness, but neither of these alter the symmetry of the image. Image simulations, such as they are generally practiced, do not tackle the main real difficulty in HREM, the effects of beam tilt and crystal tilt.

Another useful application for CRISP could be as a tool for a more quantitative evaluation of electron-optical parameters and performance of electron microscopes. The presence of one or two lattice fringes, for example, of gold at very high resolution does not show that all information inside that resolution is in fact correct or even useful. In our experience with CIP, the amplitudes fall off abruptly, significantly before the Scherzer limit of the different microscopes. Furthermore, the reflections with measurable amplitudes at the highest resolution usually have such bad phases that they cannot be reconstructed, and thus have to be deleted before calculating the density maps. A more exact characterization of electron-optical parameters of different electron microscopes calls for a quantitative analysis of amplitudes and phases over the whole resolution range out to the Scherzer limit.

Finally, there may be many applications for CRISP outside electron microscopy, for other periodically patterned objects than crystals, but this falls outside the scope of this journal.

There are of course many things that cannot be performed by the present version (1.0) of CRISP. Especially for molecular biology the most severe limitation is that the algorithms for combining series of images taken at different tilt angles into one 3D structure are not included yet. For materials science it should be pointed out that CIP is only suitable for crystalline samples. Imperfections, faults and dislocations are not periodical, and such features would be smeared out rather than enhanced by applying CIP. However, there may still be a reason to use CRISP in such cases, namely by analysing the adjacent crystalline areas. From a series of images, it would be possible to quickly sort out those that have too high phase residuals and or amplitudes with too high R_{sym} values. Not only the crystalline area, but also the defects will be impossible to interpret directly on HREM images where these figures of merit are bad.

Basically, CRISP reconstructs nearly correct amplitudes and phases, without trying to explain why they were distorted in the HREM image. A more advanced analysis of the errors in amplitudes and phases could lead to a scheme of correcting for the distortions by applying some mathematical function, which would be the inverse of the distortions. If such a function were applied in reciprocal space, i.e. on the whole Fourier transform, and not only on the exact lattice points, it should be possible to reconstruct the original HREM image, essentially free from distortions. Such an image could then be safely interpreted directly, and not only for the crystalline features, but also for defects, amorphous regions, quasicrystals et cetera. Although this lies in the distant future for CRISP, we may hope that the easy availability of quantitative experimental data on amplitudes and phases with this system will prompt theoreticians to combine theoretical and experimental image data in order to solve this complicated but interesting and important problem.

In recent years more and more attention has been drawn towards materials which are ordered but not in a traditional crystalline way. These include incommensurately modulated structures and quasicrystals. In the present version of CRISP it is not possible to extract phases and amplitudes from images of such compounds, but we are currently looking into the possibility of developing algorithms also for this.

Acknowledgements

References

[1] R. Henderson and P.N.T. Unwin, Nature 257 (1975) 28.
[2] R. Henderson, J.M. Baldwin, K.H. Downing, J. Lepault and F. Zemlin, Ultramicroscopy 19 (1986) 147.
[3] R. Henderson, J.M. Baldwin, T.A. Ceska, F. Zemlin, E. Beckman and K.H. Downing, J. Mol. Biol 213 (1990) 899.
[4] W. Kühlbrandt and D.N. Wang, Nature 350 (1991) 130.
[5] S. Hovmöller, A. Sjögren, G. Farrants, M. Sundberg and B.-O. Marinder, Nature 311 (1984) 238.
[6] D.X. Li and S. Hovmöller, J. Solid State Chem. 73 (1988) 5.
[7] D.N. Wang, S. Hovmöller, L. Kihlborg and M. Sundberg, Ultramicroscopy 25 (1988) 303.
[8] K.H. Downing, H. Meisheng, H.R. Wenk and M.A. O'Keefe, Nature 348 (1990) 525.
[9] L.A. Amos, R. Henderson and P.N.T. Unwin, Prog. Biophys. Molec. Biol. 39 (1982) 183.
[10] W.T. Holser, Z. Kristallogr. 110 (1958) 266.
[11] International Tables for X-ray Crystallography. Vol. A (Reidel, Dordrecht, 1983).
[12] W.O. Saxton, T.J. Pitt and M. Horner, Ultramicroscopy 4 (1979) 343.
[13] M. van Heel and W. Keegstra, Ultramicroscopy 7 (1981) 113.
[14] J. Frank, B. Shimkin and H. Dowse, Ultramicroscopy 6 (1981) 343.
[15] M.A. O'Keefe, P.R. Buseck and S. Ijima, Nature 274 (1978) 322.
[16] W.O. Saxton and D.J. Smith, Ultramicroscopy 18 (1985) 39.