Ultramicroscopy 25 (1988) 303–316
North-Holland, Amsterdam

STRUCTURE DETERMINATION AND CORRECTION FOR DISTORTIONS IN HREM BY CRYSTALLOGRAPHIC IMAGE PROCESSING

D.N. WANG and S. HOVMÖLLER

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

and

L. KIHLBORG and M. SUNDBERG

Department of Inorganic Chemistry, University of Stockholm, S-106 91 Stockholm, Sweden

Received 1 October 1987; received in final form 25 March 1988

A heavily distorted high resolution electron microscopy (HREM) image of a cesium–niobium oxide–fluoride sample wa reconstructed by crystallographic image processing (CIP). In the reconstructed image, showing 4bm symmetry, all metal atom were clearly seen with correct contrast, whereas in the original micrograph the fourfold symmetry was lost and only som areas could be interpreted in terms of an atomic arrangement. Defocus, astigmatism and electron beam and crystal tilt wer evaluated by CIP, and it was found that the distortion in this case was mainly due to beam tilt and crystal tilt. After correctin for the distortions, a calculated map was compared to the structure of an apparently isomorphous compound ~ $TlNb_7O_{18}$ previously solved by X-ray diffraction. The atomic positions agreed to within ±6 pm. A simulated image with electron bean tilt was in good agreement with the experimental one.

1. Introduction

Crystallographic image processing (CIP) of electron micrographs has been in use for almost twenty years in molecular biology, especially for the study of th ı crystals [1]. Over one hundred studies of protein crystals by electron microscopy and image processing have been published, many of them with three-dimensional structure determinations [2]. In molecular biology the main advantage of image processing is the noise reduction, since the noise often is stronger than the signal for both stained and unstained protein structures. Another very important reason for using image processing in molecular biology is the radiation sensitivity of biological molecules. In order to obtain high resolution or even medium resolution data (2 nm) it is necessary to collect information from many molecules, and to average over those by, for example, Fourier techniques.

For inorganic materials the situation is differ ent. Many of them are quite insensitive to radia tion damage, typically tolerating an electron dos a million times higher than that which destroys protein completely. Also the noise is less for a inorganic crystal, since it does not need a stain t improve the contrast or an embedding medium t stabilize the structure. But there are still ac vantages of image processing that can be ver fruitful for improving the quality of HREM pi tures. By quantitative analysis of the amplitude and phases that make up an image, it is possibl to correct for electron-optical distortions. Th crystallographic symmetry can be imposed, an the resulting density map can then give a rath accurate structure. In a study of $K_7Nb_{15}W_{13}O_8$ Hovmöller et al. [3] showed that atomic coord nates for the heavy metal atoms were determine with an accuracy of ±10 pm by HREM and CI by comparing the model derived from electrc

microscopy with the structure of an isomorphous compound determined by X-ray diffraction.

Smith et al. demonstrated the great importance of beam alignment in HREM [4,5]. A beam tilt of only 2 to 3 mrad could shift apparent atom positions by as much as 20 to 30 pm. They concluded that without correcting for this effect, interpretation of HREM images in terms of atomic coordinates must be made with great caution.

Recently Henderson and his coworkers [6] showed that it is possible using CIP to detect and quantify most of the electron-optical effects which may distort the images in the electron microscope, including beam tilt. This is possible because the image contains both amplitude and phase information for a large number of Fourier terms, whereas the electron-optical distortions are described by only a few variables. The crystal symmetry is, of course, very valuable in determining the distortions. By correcting for electron-optical distortions, and also some crystal distortions that are not likely to be as serious for inorganic compounds as for proteins (such as bending), Henderson et al. could determine the phases of a crystal of bacteriorhodopsin to 0.35 nm resolution. It was possible to reconstruct data both from heavily astigmatic images and from images with beam tilt large enough to cause phase errors of over 360°.

In the present investigation we have applied CIP to high resolution (0.25 nm) images of a cesium–niobium oxide–fluoride sample of unknown composition. The electron diffraction patterns of this phase indicated tetragonal symmetry with $a = 2.76$ nm and $c = 0.39$ nm. These observations together with several general features of the HREM images recorded were compatible with a structure of "$Rb_3Nb_{54}O_{146}$" type [7]. However most of the images did not show tetragonal sym-

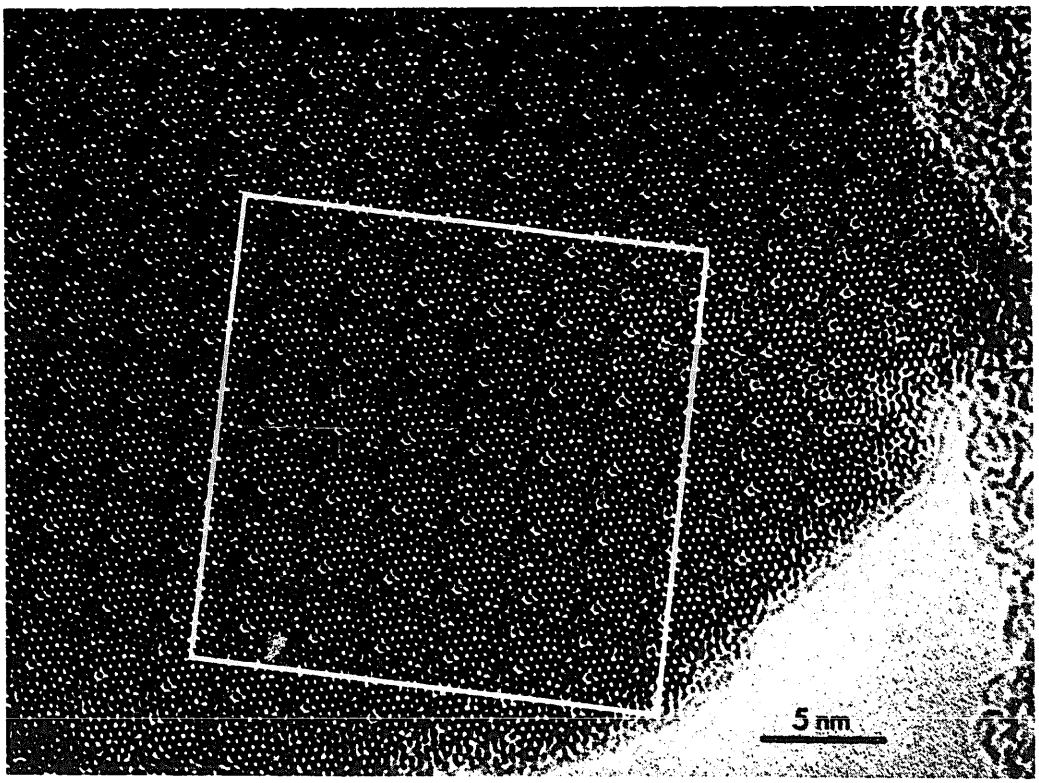

Fig. 1. HREM image of $Cs_xNb_{54}(O,F)_{146}$. Although the electron diffraction pattern shows 4mm symmetry, this image has no 4-fol symmetry. Notice that in some areas in each unit cell atom-like features are white, while in other areas similar features appear a black points. This appearance is found in every unit cell of the area in the picture where the crystal is thin. An amorphous region lik the one at the upper right corner is especially suitable for evaluating defocus and astigmatism of the picture, by optica diffractometry. The boxed area was digitized and processed. Electron optical magnification 500,000×.

metry. This was particularly true for one set of images with excellent resolution taken from a very thin part of a crystal fragment. In one direction the contrast features could easily be interpreted in terms of an atomic arrangement, while in the perpendicular direction the interpretation was not straightforward. In order to determine the structure unequivocally and to find the reason for the deviation from tetragonal symmetry in the micrographs but not in the electron diffraction pattern, one of the images, that which seemed to be recorded close to Scherzer focus (fig. 1), was used for image reconstruction by CIP. We concluded that there must be enough redundant information in the image to determine the effects that distorted the picture, and we expected to be able to solve the structure by finding and correcting for these distortions. After the structure had been solved from this single electron micrograph, we applied CIP to two more images, recorded from another fragment, and obtained similar structural results.

2. Electron microscopy

A few large colourless crystals selected from a multiphase cesium–niobium oxide–fluoride sample were crushed in an agate mortar and dispersed in *n*-butanol. Drops of the resultant suspension were collected on a perforated carbon film supported on a Cu grid. The grid was then examined in a JEOL 200CX electron microscope, equipped with an ultrahigh resolution top-entry goniometer stage with tilt angles of $\pm 10°$ and operated at 200 kV. The radius of the objective aperture used corresponded to 4.1 nm^{-1} in reciprocal space, and the HREM technique was applied. Thin crystal fragments projected over holes in the carbon film were aligned with the short (~ 0.39 nm) crystal axis parallel to the electron beam.

3. Image processing

The electron micrograph (fig. 1) was first checked in an optical diffractometer. The optical diffraction pattern had sharp and strong diffraction points out to 11 orders or more, corresponding to at least 0.25 nm resolution. Different areas of the micrograph were compared, and the optical diffraction pattern was fairly uniform over the thin part of the crystal. In a region of amorphous material the shape of the contrast transfer function (CTF) could be seen to be practically circularly symmetric, indicating only little astigmatism. The intensities of the spots in the optical diffraction pattern indicated 4mm symmetry. A region of about 50 unit cells near the edge of the crystal (indicated in fig. 1) was chosen for further processing, on the basis of being a very thin part with strong and symmetric diffraction to high resolution. Due to the very limited thickness, the weak phase object approximation should hold reasonably well.

The selected area was scanned in a Joyce–Loebl flat-bed microdensitometer MDM 6, using a raster size of 40×40 μm, and recording 256 by 256 raster points. The total area scanned was then about 1 cm^2. A raster size of 40 μm at 500,000 times magnification corresponds to a sampling size of 0.08 nm in the crystal, fine enough to preserve information to at least 0.25 nm resolution.

The digitized image was transferred to a VAX 11/750 computer, where all further data processing was done. The Fourier transform of the scanned image was calculated. This gives both the amplitude and the phase information. An image of a periodic object can be described as a sum of cosine waves of different periodicities (i.e. wavelengths), denoted by the (h, k) indices, each with a specific height (amplitude) and position (phase relative to a given origin point in the image. Initially the origin is simply the center of the scanned area, a point which usually does not coincide with a symmetry element.

The reciprocal lattice was indexed on the computer graphics display, and later refined using the peak coordinates of 12 strong reflections. Based on this accurate lattice, the amplitudes and phases for all reflections out to 0.25 nm resolution were extracted from the Fourier transform. Amplitudes were calculated by integrating over the 3×3 grid points closest to the predicted position of the

lattice point in the Fourier transform, followed by subtraction of the local background. The phases were taken as the phase value at that grid point in the calculated Fourier transform which was closest to the predicted position of the lattice point.

The positions of the expected 4-fold axes in the unit cell were found by calculating the phase residual modulo 180° at every position in the unit cell, as described earlier [3]. When the origin is on a 2- or 4-fold axis all phases should be 0° or 180°. The initial, arbitrary, origin of the unit cell was shifted to coincide with one of the 4-folds by shifting the phase values of all reflections by $360(h \cdot x + k \cdot y)$ degrees, where (x, y) is the position of a 4-fold axis in the unit cell relative to the initial, arbitrary origin.

A Fourier transform of a two-dimensional function is a two-dimensional array of complex numbers $A + iB$. The amplitude part of the Fourier transform corresponds to the ED pattern (fig. 2a) and to the optical diffraction pattern, with the important difference that the ED pattern, unlike the other two, has not been attenuated by the CTF. In addition, the ED pattern relates to a much larger area of the crystal than that scanned. The computer-calculated diffraction pattern was displayed on a Digital VS11 raster graphics colour screen (fig. 2b).

ED intensities were determined by integrating over the diffraction spots after scanning the ED negative in the microdensitometer.

4. Image simulations

From structure models derived, images were calculated using a locally modified version of O'Keefe's SHRLI suite of programs [8]. The development of the wave field within the crystal was approximated by the multislice method [9] and the imaging system was simulated by an aberrated lens. The insertion of patches of the calculated images into the observed one was made on a KONTRON IPS image processing system.

5. Determination of defocus and astigmatism

An objective lens in an electron microscope does not give an undistorted enlarged image of the object. The image formation is affected by several distorting factors inherently present in the microscope. These effects influence both amplitudes and phases of the diffracted beams. Consequently it may not always be justified to interpret an electron micrograph directly. However, the differ

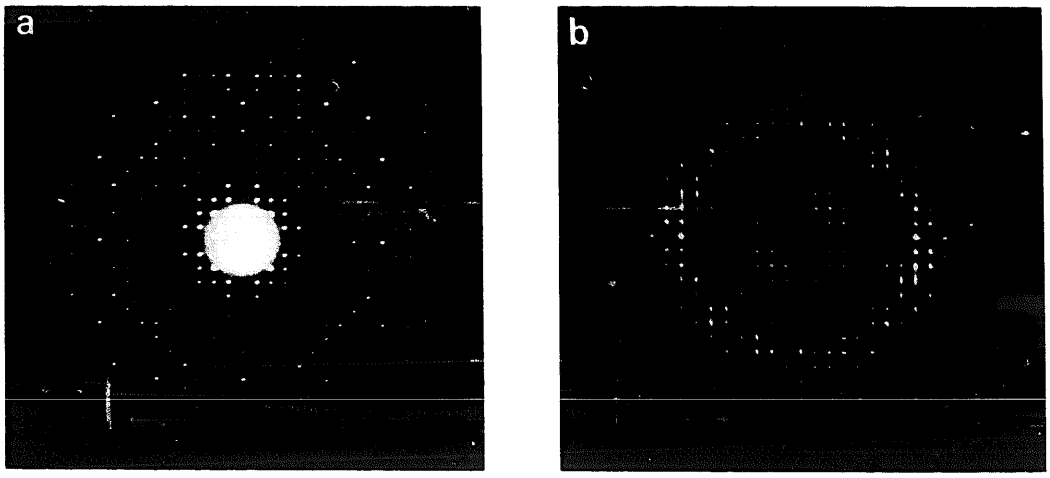

Fig. 2. (a) Electron diffraction pattern from the crystal shown in fig. 1. The symmetry is tetragonal 4mm, and it diffracts to some 1 orders, corresponding to 0.16 nm. (b) Computer-calculated diffraction pattern from the boxed area in fig. 1 shows the same symmetr as the electron diffraction, and it extends equally far in all directions. The calculated diffraction extends only to about 11 orders, c 0.25 nm resolution, due to the damping effects in the electron microscope.

ent distortions can be detected and quantified by CIP, and then corrected for, with a more or less distortion-free image as a final result.

The number of parameters that determine the distortions in the image is very limited. For a thin inorganic crystal one only needs to take into account the following seven parameters:

defocus	1 parameter,
astigmatism	2 parameters,
beam tilt	2 parameters,
crystal tilt	2 parameters.

The other parameters like spherical aberration C_s, chromatic aberration C_a, beam convergence and the size of the aperture used must also be taken into account, but these can be treated as either known constants of the instrument (C_s and C_a) or factors which limit the resolution by damping the CTF at high resolution. The image formed is then a function of a limited number of unknown parameters and the noise, as:

structure + crystal orientation

+ electron-optical effects + noise → image.

From our earlier study of the compound $K_7Nb_{15}W_{13}O_{80}$ we know that the noise level of HREM images of such metal oxides, taken over holes in the holey-carbon film, can be very low [3]. The problem of reconstructing an undistorted image is then reduced to finding and compensating for the 7 parameters listed above. Unfortunately these parameters give rise to nonlinear effects, but it is still possible to estimate the parameters quantitatively. The problem may be thought of as finding the solution to a system of equations with 7 unknown parameters and about 100 equations, namely the amplitudes and phases of the diffraction points of the image.

Fortunately the different parameters influence the image in different ways, and they can therefore be determined one by one at least to a first approximation. The effect of the defocus is perhaps the most well known. It changes the appearance of the CTF. At Scherzer focus all reflections within the resolution limit of the microscope are recorded with the correct phase. Only the amplitudes are distorted, namely by being attenuated by the smooth CTF. Outside the Scherzer limit the phases are alternately wrong or correct. If the image is not recorded at Scherzer focus, the CTF changes sign already at a lower resolution. When the CTF changes sign, the reflections change phase by exactly 180°, i.e. they are reversed in contrast, making black white and vice versa. Further out in resolution the CTF changes sign again and the phases in this region will be correct. The CTF continues to oscillate in this manner, giving rise to rings of reflections with phase errors of 0° 180°, 0°, 180°, ..., until finally the CTF is damped out by beam convergence and focus spread and all reflections vanish.

When an image is astigmatic, the imaging system is not cylindrically symmetric. This is equivalent to having different defocus values in different directions of the image. To a good approximation the astigmatism can be described by only two parameters; the direction of maximum defocus and the magnitude of the difference in focus between the average and maximum defocus values. The effect of astigmatism is to change the above mentioned rings of alternating contrast caused by the CTF into ellipses. Just as was the case for defocus, image phases are changed only in multiples of 180° due to astigmatism. If an image is affected by astigmatism this can be seen in the optical diffraction pattern. An amorphous region is especially useful for this purpose, owing to the continuous character of its diffraction pattern where the elliptical appearance of the CTF is easily seen. To a good approximation the astigmatism and focus are constant over the whole area recorded in one micrograph, and it is possible to determine the astigmatism from a small amorphous region outside the crystal. In the calculated Fourier transform astigmatism can also be determined by comparing amplitudes and phases of symmetry-related reflections. Symmetry-related reflections have the same vector length in Fourier space, and they have the same amplitudes. In an astigmatic image this symmetry is lost, and it can be verified that it is really an effect of astigmatism by the sudden phase changes of 180° for reflections just outside a node of the CTF, i.e. outside the point of reversal of sign for the CTF. It is important to remember that neither defocus nor astigmatism

will cause a gradual change in phase values, but only sudden shifts of 180°.

6. Effect of beam tilt and crystal tilt

If the electron beam is not running down exactly along the optical axis of the electron microscope, the beam tilt will cause distortions of the image. These distortions can be distinguished from those caused by defocus and astigmatism, in that the beam tilt for a thin specimen to a first approximation only affects the phases, and that the phases are changed by [4]:

$$\Delta\phi = -2\pi\lambda\left(C_s\lambda^2K^2 - \mathrm{d}f\right)\boldsymbol{K}_0\cdot\boldsymbol{K},$$

where $\boldsymbol{K}$ is a reciprocal vector and $\boldsymbol{K}_0$ the beam tilt vector, which represents the direction and the magnitude of deviation of the electron beam from the optical axis. In one direction of reciprocal space, along the direction of maximum beam tilt, the phases will change most significantly, but in the perpendicular direction the phases will be unaffected. It is clear that both astigmatism and beam tilt can cause an image to lose the true symmetry of the crystal, but by CIP it is possible to distinguish between these effects in pictures.

The effects of crystal tilt are somewhat similar to those of beam tilt, but for thin crystals they are less severe for the same magnitude of tilt [4]. There is a possibility that the crystal is bent, and that the thin edge has a different orientation from the main part of the crystal. This can be tested by inspecting the optical diffraction pattern from the small area that is to be scanned. If this area of the crystal is tilted, then the diffraction pattern will become asymmetric in intensity, with a loss of high resolution reflections in the direction perpendicular to the tilt axis. In thicker parts of the crystal these effects become even stronger.

7. Results

It is clear from the electron micrograph (fig. 1) that the 4-fold symmetry has been lost in the image. In some parts of the unit cell black dots are seen on a light background, while in other parts the contrast seems reversed. Thus the image must have been distorted by crystal tilt, beam tilt and/or astigmatism.

We first evaluated the possible astigmatism of the image. There was an area of amorphous material adjacent to the crystal (fig. 1). The optical diffraction pattern of this area was fairly symmetrically round indicating low astigmatism. We also estimated the astigmatism by comparing the amplitude obtained from the Fourier transform of the image with the intensities obtained by ED. The amplitudes of all 160 reflections in the unique half of the diffraction pattern are listed in table 1. The CTF in different directions of reciprocal space i.e. the ratios between amplitudes obtained from the Fourier transform of the image and from the electron diffraction pattern, were plotted as functions of reciprocal vector length along different directions. Two of them, along [100] and [010] respectively, are shown in figs. 3a and 3b. The CTF was found to be similar in the different directions and thus the astigmatism is quite small.

The defocus value can also be determined from fig. 3, since different defocus values will change the CTF and thus the attenuation of reflections at different resolution. From the ED intensities we can only obtain the square of the CTF, but it is still possible to distinguish the two possible cases of defocus values $\mathrm{d}f = -100$ nm and $\mathrm{d}f = -67.5$ nm. The contrast transfer functions are shown in figs. 3c and 3d for these two defocus values together with the experimentally determined ratios between the amplitudes in the image and in the electron diffraction pattern. The amplitude ratios illustrated in figs. 3a and 3b will follow the absolute values of the CTF. For the defocus value of -100 nm (fig. 3c) the magnitudes of the CTF will be close to zero around the cross-over point at about 2 nm^{-1}, while for -67.5 nm defocus the magnitudes will have a large value at 2 nm^{-1} and a local minimum near 3 nm^{-1}. The ratios of the observed amplitudes indicate a defocus value of about -100 nm, since they are close to zero around 2 nm^{-1}. Also, a ring of weak reflection around the 5th-order spots in the Fourier transform supports this conclusion (fig. 2a). Consequently, the originally observed phases of all the

Table 1
Observed (after CTF correction) and deduced phases and amplitudes for the Cs–Nb–O–F compound; the observed phases and amplitudes were extracted from the Fourier transform of the negative of the electron micrograph shown in fig. 1; the deduced phases were arrived at by imposing exact P4bm symmetry, and minimizing difference of phase error between adjacent diffraction spots in reciprocal space; the amplitudes were averaged over the 4 (or 2) symmetry-related diffraction points. The deduced phases and averaged amplitudes were used to calculate the corrected density map (fig. 5)

h k	Observed phases (deg)				Deduced phases (deg)				Observed amplitudes				Averaged amplitudes
	(h, k)	$(h, \bar{k})$	(k, h)	$(k, \bar{h})$	(h, k)	$(h, \bar{k})$	(k, h)	$(k, \bar{h})$	(h, k)	$(h, \bar{k})$	(k, h)	$(k, \bar{h})$	
1 1	147	161	–	–	180	180	–	–	1020	1459	–	–	1239
2 0	150	–	–	195	180	–	–	180	2399	–	–	1987	2193
2 1	158	327	348	182	180	0	0	180	2378	2270	2495	2270	2353
2 2	338	337	–	–	0	0	–	–	2891	2763	–	–	2827
3 1	328	330	350	353	0	0	0	0	1558	1697	1921	1881	1764
3 2	134	308	345	163	180	0	0	180	1533	1760	1578	1682	1638
3 3	28	85	–	–	0	0	–	–	512	719	–	–	615
4 0	316	–	–	324	0	–	–	0	648	–	–	506	577
4 1	332	303	172	89	0	180	180	0	674	425	780	1056	733
4 2	332	338	344	286	0	0	0	0	996	612	942	581	782
4 3	139	327	336	155	180	0	0	180	88	736	267	582	418
4 4	149	145	–	–	0	0	–	–	1429	792	–	–	1110
5 1	125	193	201	91	0	0	0	0	309	491	368	407	393
5 2	304	137	154	301	180	0	0	180	692	1130	705	1135	915
5 3	152	302	143	198	0	0	0	0	556	282	501	−72	316
5 4	348	153	141	246	180	0	0	180	683	675	717	520	648
5 5	152	128	–	–	0	0	–	–	2654	1272	–	–	1963
6 0	185	–	–	233	0	–	–	0	162	–	–	−70	46
6 1	218	281	78	110	0	180	180	0	264	343	431	210	312
6 2	118	143	188	87	0	0	0	0	1239	1175	1427	1465	1326
6 3	350	257	169	132	180	0	0	180	311	59	421	−296	123
6 4	122	106	161	109	0	0	0	0	845	745	975	804	842
6 5	318	116	154	255	180	0	0	180	2431	1564	2593	1816	2101
6 6	333	268	–	–	180	180	–	–	5456	3098	–	–	4277
7 1	160	169	242	49	0	0	0	0	496	755	889	1099	809
7 2	275	132	206	236	180	0	0	180	648	974	625	1278	881
7 3	122	147	199	56	0	0	0	0	2680	3018	4156	4196	3512
7 4	136	319	5	41	0	180	180	0	1752	1908	2466	2347	2118
7 5	310	328	344	231	180	180	180	180	441	469	1046	731	671
7 6	163	312	352	67	0	180	180	0	1706	1316	2134	1265	1605
7 7	162	264	–	–	0	0	–	–	692	452	–	–	572
8 0	178	–	–	60	0	–	–	0	3318	–	–	5252	4285
8 1	185	359	103	36	0	180	180	0	1038	2056	2389	3046	2132
8 2	344	348	73	204	180	180	180	180	1950	3629	4205	5210	3748
8 3	133	339	47	19	0	180	180	0	2047	3267	4392	4894	3650
8 4	336	296	12	84	180	180	180	180	183	165	306	379	258
8 5	156	326	17	7	0	180	180	0	1909	2108	3231	2617	2466
9 1	208	198	317	360	0	0	0	0	1467	1930	3024	3200	2405
9 2	54	199	302	164	180	0	0	180	821	1675	2393	2906	1948
9 3	90	66	93	158	0	0	0	0	251	328	285	217	270
9 4	257	25	71	327	0	180	180	0	619	1211	1746	1771	1336
9 5	74	6	38	151	180	180	180	180	242	592	737	834	601
10 0	66	–	–	149	0	–	–	0	1335	–	–	2888	2112
10 1	201	7	174	17	180	0	0	180	352	426	519	521	454
10 2	180	224	20	271	180	180	180	180	168	150	391	293	250
10 3	240	329	129	23	180	0	0	180	161	264	83	−217	72

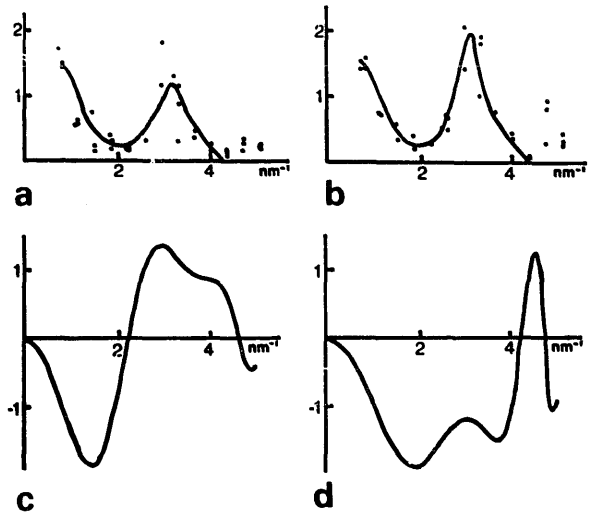

Fig. 3. Determination of defocus and astigmatism. The amplitudes obtained from the Fourier transform of the image were divided by those from electron diffraction and the ratios were plotted as a function of reciprocal vector length in different directions, giving a two-dimensional function of |CTF|. All the curves are similar, proving that the astigmatism is negligible. Two of them, along the [100] and [010] directions respectively, are shown in (a) and (b). The defocus value of the image was determined by comparing the functions in (a) and (b) to the absolute value of the CTF. The |CTF| for JEOL 200CX at $df = -100$ nm (c) and $df = -67.5$ nm (d) are similar to the experimental curves. For any other defocus value the |CTF| is not in close agreement with the curves in (a) and (b). The defocus of -67.5 nm is discarded because the amplitude of its CTF has a large value at 2 nm^{-1}, where the amplitudes of the experimental curves are close to zero. Therefore, the defocus value of the image is determined as -100 nm.

reflections outside 2 nm^{-1} frequency have been reversed for further investigations, as shown in the first columns in table 1.

The image was distorted by crystal tilt. This could be seen already on the optical diffractometer. In the thin area of the crystal the diffraction pattern is quite symmetric, but in the thicker areas of the crystal higher-order reflections along the direction (4, -1) decreased in intensity and eventually vanished. A quantitative estimate of the crystal tilt could be made by comparing amplitudes of symmetry-related reflections (see table 1). For the reflections (7, 1), (7, 2), (7, 3), (8, 1), (8, 2), (8, 3) and (9, 1), (9, 2), (9, 3) the amplitudes of symmetry-related reflections decrease in the order of $(k, -h) > (k, h) > (h, -k) > (h, k)$. This systematic relation is very striking and can definitely not be attributed to random noise. The amplitudes of $(k, -h)$ are 2 to 3 times larger than those of (h, k). For a crystal thickness of 3.94 nm (10 slices) a crystal tilt of (60, -15), equivalent to 56 milliradians, will simulate the amplitude values of table 1 very well.

8. Space group determination

The phases extracted from the Fourier transform after shifting the origin to a 4-fold position are seen in table 1. A projection down a 4-fold axis is centrosymmetric, and therefore all phases should be 0° or 180°. The phases have been listed in groups of the four symmetry-related reflections. The 4-fold symmetry causes the two reflections (h, k) and $(-k, h)$ to have identical amplitudes and identical phases. The relation between a pair of reflections (h, k) and $(h, -k)$ is not quite so straightforward. The mirror-symmetry of the electron diffraction pattern along the axes and diagonals can be caused by any of four different symmetry elements: a mirror plane, a glide plane, a 2-fold rotation or a 2_1 screw axis. In projection it is not possible to distinguish between a mirror plane and a 2-fold axis. Similarly a glide plane and a 2_1 screw axis are indistinguishable in projection. However, if the phases are known, a mirror plane or 2-fold axis can be distinguished from a glide plane or a 2_1 screw. In all four possible cases the reflections (h, k) and $(h, -k)$ will have equal amplitudes. For mirror planes and 2-fold axes the phases will be identical within such a pair of reflections, whereas for a glide plane or a 2_1 screw axis the reflections for which $(h + k)$ is an odd number will differ in phase by 180°, and only reflections with $(h + k)$ even will have equal phases.

The electron diffraction pattern alone is not sufficient for making this space-group determination, unless clear systematic absences along the crystal axes can be seen. The electron diffraction pattern in this case is not conclusive. Even reflections are much stronger than the odd ones along the crystal axes, but there are some weak reflections with odd indices. If such reflections are found in X-ray diffraction it proves that there are

no systematic absences, but in electron diffraction space-group-forbidden reflections may well arise from dynamical scattering.

The space-group determination is preferentially done by comparing phases of symmetry-related reflections. For high-order reflections (h and/or k larger than 4), the phase relationships are hard to see (table 1), but for lower frequencies the phases are clearly distinguished as being close to either 0° or 180°. It is evident by looking at the two pairs (2, 1), (2, −1) and (3, 2), (3, −2) that the phases differ by 180° for $(h+k)$ odd, and therefore the symmetry must be a glide plane or a 2_1 screw along the axes. The possible space groups are then P4bm and $P42_12$, these two space groups being indistinguishable in projection.

9. Phase determination

For the space groups P4bm and $P42_12$ every group of four symmetry-related reflections only has two possible combinations of phases. For reflections with $(h+k)$ even, all four reflections

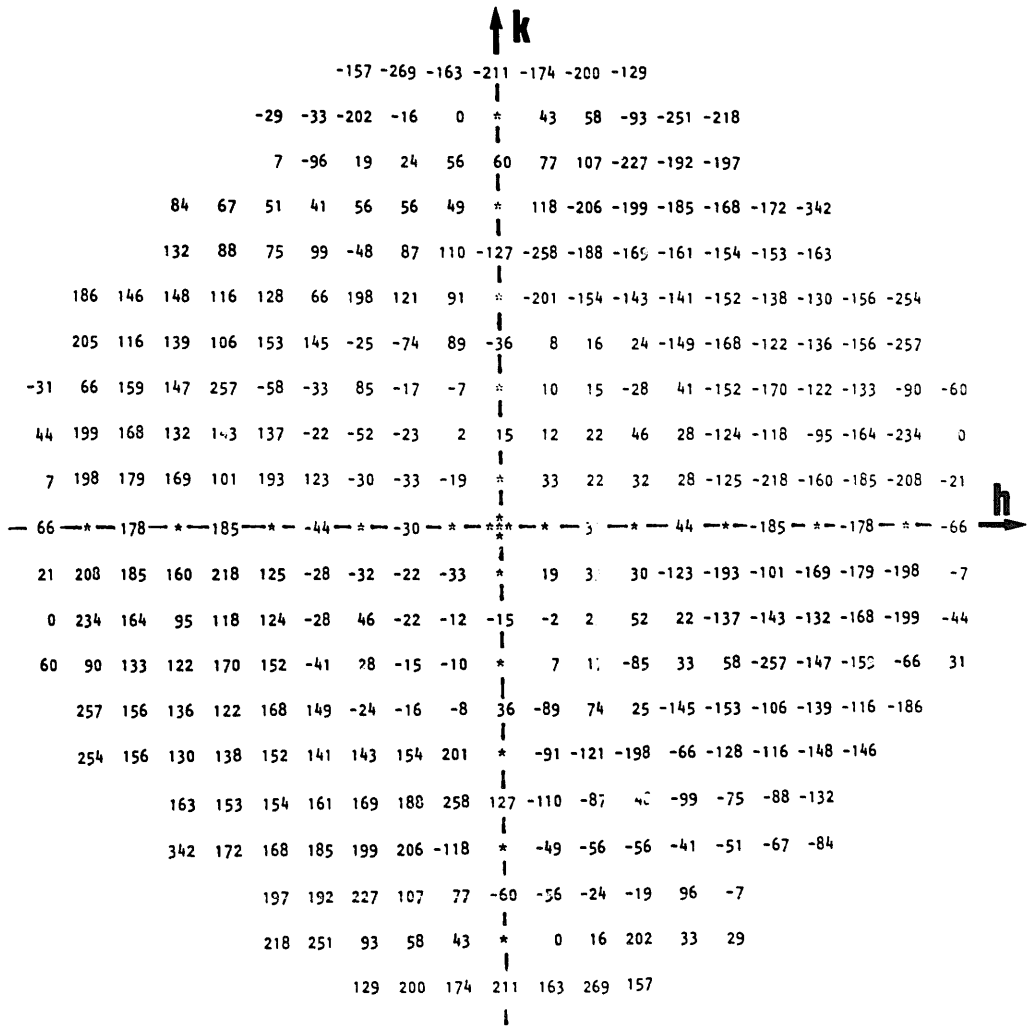

Fig. 4. Phase errors plotted as a function of position in reciprocal space, i.e. (h, k) index. The values are the differences between tl corrected and the observed phases (after CTF correction). Systematically absent reflections ($(h, 0)$ with h odd and $(0, k)$ with k od are marked by *. The phase errors are relatively small for low-order diffraction points (h and $k \leqslant 4$), and also for the diffracti points close to the direction [−110], while for the diffraction points along [110] the phase errors increase dramatically with resolutio implying a beam tilt in this direction in the electron microscope.

must have a phase of 0° or they must all be 180°. For reflections with $(h + k)$ odd, either (h, k) and $(k, -h)$ are 0° with $(h, -k)$ and (k, h) being 180°, or vice versa. The four symmetry-related reflections have been extracted independently out of the Fourier transform so that we have four independent estimates of the phases for each quartet of reflections. If, for example, three reflection are reliably estimated, and one is deviating, it can be set to obey the others. Because the phase errors will only change slowly in reciprocal space, two reflections that are close in the diffraction pattern, for example (7, 4) and (8, 3), will be distorted in more or less the same way. The unpredictable noise is not expected to cause phase errors of more than 5° or 10°, which is not large enough to create any serious problems for assigning reflections a phase of either 0° or 180°.

The phases of all reflections for this Cs-Nb-O-F compound were solved in the following way. First all reflections out to (4, 0) were given phases directly, since they are all very close to 0° or 180° and obey the phase relation rules for this symmetry. The reflection (4, 1) did not fit well to any of its two possible sets of phase values, so it could not be determined with certainty. The quartets (4, 2) and (4, 3) can easily be given values, as seen in table 1. With the phase combination suggested in table 1 for these two quartets the total phase errors are 140° and 123° respectively, but if the phase would have been reversed the phase errors would be 580° $(= 4 \times 180° - 140°)$ and 597°, respectively. Moreover, the phase errors for two reflections that are close in reciprocal space should be similar in sign and magnitude, and all reflections should therefore have their phase increased by between 16° and 74°.

With application of this procedure the phases are gradually solved one by one, always minimizing the difference in phase errors between adjacent reflections, until only a few reflections are left undetermined. These could be set to amplitude zero. Examples of such difficult reflections are the (4, 1) mentioned above and (6, 1) and (6, 3). These difficulties are due to the cross-over of the CTF at 2 nm^{-1}. Near the cross-over the amplitudes of the reflections are much reduced, and the phases become more difficult to determine. Also if there is a slight astigmatism two symmetry-related reflections may fall on either side of the cross-over, in which case the one outside has been reversed in phase relative to the one inside. Outside 0.38 nm^{-1}, the phases were hard to determine, and the data beyond this limit have not been included.

After the phases have been solved in this way, we have a list of finally-corrected phases, as shown in the middle columns of table 1. The differences between these phases and the observed phases (but after CTF correction) were plotted as a function of indices (h, k) in reciprocal space (fig. 4). From this figure it was observed that the phases were close to correct (within ±30°) only in the low resolution area. In the direction $(h \sim k)$, the phase errors increase dramatically with resolution. This is a clear indication of electron beam tilt. The fact that the phase errors of high-order reflections

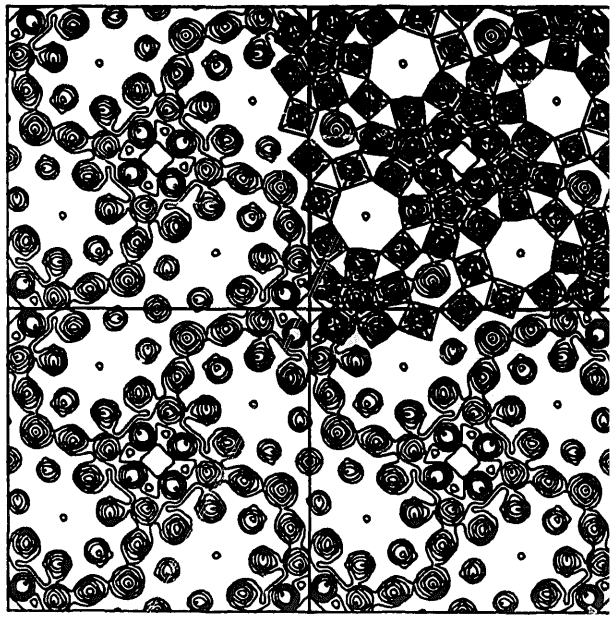

Fig. 5. Density map of the $Cs_xNb_{54}(O,F)_{146}$ compound after applying crystallographic image processing (CIP). The corrected phases and averaged amplitudes listed in table 1 were used as input. The characteristic TTB block of NbX_6 octahedra can be seen around the 4-fold axes. Over the upper-right unit cell is overlaid the model of $TlNb_7O_{18}$ determined by X-ray diffraction (after ref. [10]). The thallium atoms are not drawn in the X-ray model, but they are found inside the six- and seven-sided tunnels formed by NbX_6 octahedra, with higher density in the six-sided tunnels. The occupancy of Tl atoms in the X-ray investigation was found to be 0.64 for the six-sided and 0.46 for the seven-sided tunnels.

change sign relative to the low-order reflections, is also an intimation of beam tilt. The direction of beam tilt must be close to the diagonal $h = k$, where the phase errors are largest.

10. The structure map

The potential map of the structure is calculated as the Fourier transform of the corrected list of amplitudes and phases (table 1). Every general reflection occurs 8 times in reciprocal space as symmetry-related reflections. All symmetry-related reflections were given the same amplitudes, namely those of the average of those reflections. We tried to use both the amplitudes derived from the Fourier transform of the image and amplitudes obtained by measuring the intensities in the electron diffraction pattern. Although the latter do not suffer from the nonuniform attenuation caused by the CTF, it gave a poorer density map. The reason for this is that the ED comes from a large part of the crystal with variable thickness and orientation.

The calculated density map (fig. 5) shows peaks that could be attributed to niobium and cesium atoms. As initially suspected, the structure is isotypic with that of "$Rb_3Nb_{54}O_{146}$" [7] and ~ $TlNb_7O_{18}$ [10]. The structure is characterized by TTB (tetragonal tungsten bronze) units, where all five-sided tunnels are occupied by –Nb–X–Nb–X– strings (X = O, F) so that four pentagonal columns (PCs) [11] are formed. All PCs have one NbX_6 octahedron in common with the next neighbouring PC. The TTB units are mutually linked via additional NbX_6 octahedra forming six- and seven-sided tunnels where the Cs atoms are located. The stoichiometry of the structure model is $Cs_xNb_{54}(O, F)_{146}$, $x \leqslant 8$. It should be emphasized that in the present case this metal atom arrangement was arrived at without using any prior knowledge of the chemistry or structure of the crystal. Only completely general crystallographic and electron optical techniques were applied to get the correct set of phases.

The structure of ~ $TlNb_7O_{18}$ which has been solved by X-ray diffraction [10] has space group P4/mbm with $a = b = 2.75$ nm, $c = 0.394$ nm. A projection along the short (~ 0.4 nm) axis gives a structure without metal atoms overlapping. Thus the main feature of the structure can be solved by this single projection. The atomic coordinates of the metal atoms reported from this phase (table 2) were compared to those of the present model measured as the centers of the peaks in the calculated potential map (fig. 5). The average difference in atomic coordinates between these two structures was only 6 pm for the 8 niobium atoms. This value is even better than the 10 pm reported for $K_7Nb_{15}W_{13}O_{80}$ [3] and 13 pm found in $Na_3Nb_{12}O_{31}F$ [12]. Both the thallium and the

Table 2
Atomic coordinates for the heavy metal atoms in the $Cs_xNb_{54}(O,F)_{146}$ compound as determined by HREM and CIP, and ~ $TlNb_7O_{18}$ as determined by X-ray diffraction [10], $a = b = 2.75$ nm; the thallium atoms were reported to be disordered inside the six- and seven-sided tunnels; thus no comparisons between Cs and Tl atom positions are given in the table

Atom	Fractional atomic coordinates				Difference EM/X-ray		Difference (Å)
	Electron microscopy		X-ray diffraction		x/a	y/b	
	x/a	y/b	x/a	y/b			
Nb (1)	1/2	0	1/2	0	0	0	0
Nb (2)	0.199	0.699	0.1986	0.6986	0.0004	0.0004	0.01
Nb (3)	0.233	0.939	0.2328	0.9379	0.0002	0.0011	0.03
Nb (4)	0.366	0.968	0.3678	0.9642	0.0018	0.0038	0.12
Nb (5)	0.064	0.720	0.0664	0.7212	0.0024	0.0012	0.07
Nb (6)	0.158	0.828	0.1583	0.8260	0.0003	0.0020	0.06
Nb (7)	0.040	0.842	0.0382	0.8390	0.0018	0.0030	0.10
Nb (8)	0.085	0.951	0.0864	0.9484	0.0014	0.0026	0.08
Cs (1)	0.094	0.594	–	–			
Cs (2)	0.311	0.311	–	–			

cesium atoms were found inside the six- and seven-sided rings of octahedra. However, in the ~ $TlNb_7O_{18}$ study [10] the thallium atoms are reported to be disordered over several close positions, making a numerical comparison of the coordinates of Tl and Cs atoms impossible.

11. Image simulations

A number of theoretical image calculations were made with the aim of simulating the original image. This work started as soon as the first guess at the structure had been made and before CIP had been attempted. Although it was suspected that beam tilt had to be introduced in the calculations to account for the loss of 4-fold symmetry in the image, a reasonable agreement was not obtained, due to the number of free parameters involved. Not until the image processing described above had been performed and an indication of the beam tilt direction had been obtained was it possible to proceed. The coordinates of all atoms were taken from [10], except for the cesium positions. These latter were assumed to be located exactly in the centres of the tunnels, with full occupancies in the six-sided tunnels and occupancies varying from 0 to 1 in the seven-sided tunnels.

The relative importance of the distortions in phases and amplitudes was evaluated in the following way. A density map was calculated using the original image phases but the 4-fold symmetry-averaged amplitudes. This map was very simi-

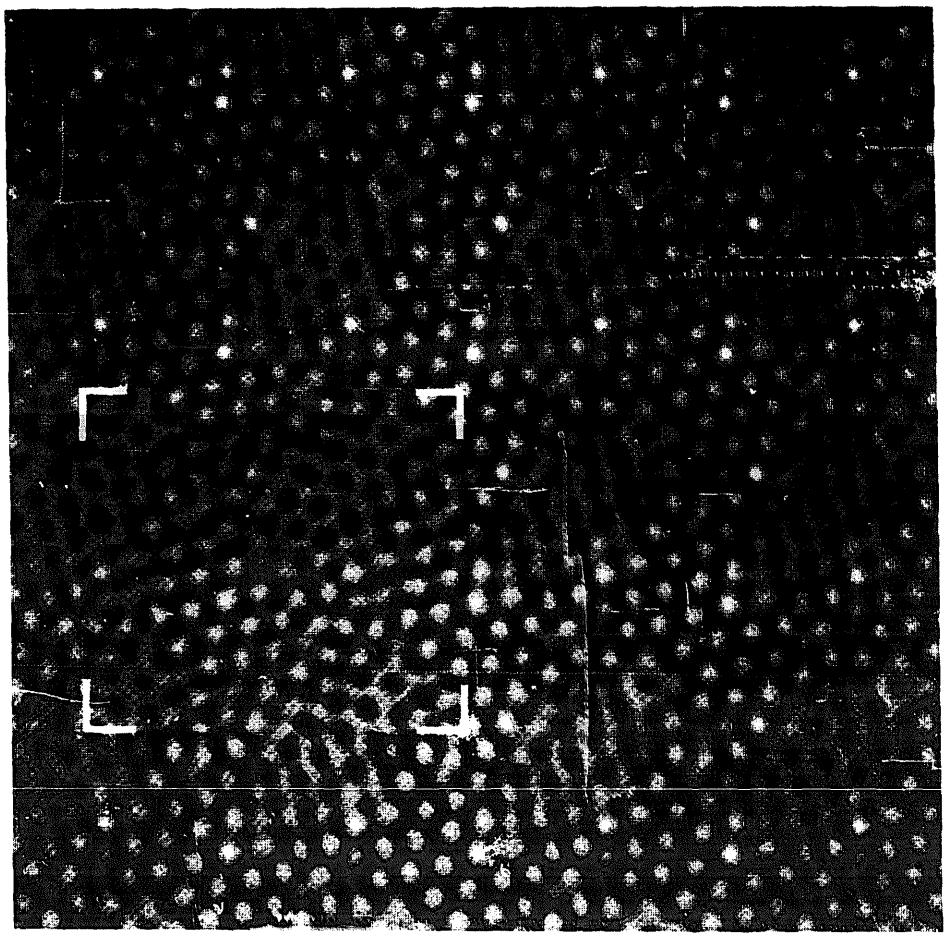

Fig. 6. Simulated image of the $Cs_xNb_{54}(O,F)_{146}$ compound inset into the original, only Fourier-filtered image. The occupancy of C in the center of the seven-sided tunnels is set to 3/4. Defocus is −103 nm, and beam tilt is 5 mrad in the [110] direction, and with n astigmatism and crystal tilt. The other imaging parameters are: high voltage 200 kV, objective aperture 4.1 nm^{-1}, spheric aberration 1.2 mm, beam divergence 5 mrad.

r to the original micrograph. However, a density ıap calculated from the original amplitudes but ith phases corrected by CIP closely resembled ıe correct structure. Therefore the distortions in ıe image were mainly caused by phase errors. To first approximation crystal tilt affects only amlitudes and electron beam tilt only phases. The istortions in this image were caused by both rystal tilt and beam tilt, but beam tilt effects were ominating.

We compared the observed and calculated nages both in real and in reciprocal space. The ıtter was possible since the image analysis as resented in this paper provides amplitudes and hases, and these data can also be obtained from he multi-slice image simulations. In reciprocal pace we could obtain a very good agreement etween the amplitudes of the observed and imulated images, using the crystal tilt with the rimary beam going along the direction of 60, −15).

In real space it was found that beam tilt magniudes larger than approximately 4.5 mrad were ıecessary for the marked asymmetry in the black/ vhite dot pattern to be correctly reproduced, and ɔnly tilt directions within ±10° from the (1, 1) liagonal and defocus values around −100 nm ;ave reasonable agreements. Fig. 6 shows the Fourier-filtered, but otherwise unprocessed, origiıal image into which a patch of a calculated image ıas been inserted. This latter is based on a multiılice calculation for 1.5 nm thickness. It is seen hat, although the black/white hexagonal dot ɔatches match reasonably well, the agreement at he seven-sided tunnels is not so good. This could ɔe due to a wrong assumption concerning the ɔesium content of the tunnel sites. The image ɔontrast was, however, rather insensitive even to large variations in the occupancy of these sites. It is also possible that the cesium atoms are not located at the very centers of these tunnels.

12. Discussion

We have shown that it is possible by using CIP to interpret HREM images of thin crystals, even when taken under suboptimum optical conditions, without using any chemical or structural knowledge. CIP differs in this respect from the image simulation methods. If both the structure and the optical conditions are unknown, a trial-and-error-type procedure can be very tedious. By CIP it is possible to determine the correct phases of the structure factors (reflections) and improve the quality of the amplitudes. From this set of virtually distortion-free structure factors a single, correct density map can be calculated by inverse Fourier transformation. By careful analysis of the deviations from symmetry in the Fourier transform of a HREM image it is furthermore possible to objectively determine the optical parameters one by one.

In this study we have shown that a structure can be solved even from a single, distorted image, provided the resolution is high enough. Usually one has several images of the same structure, and this makes CIP even easier and safer. The ambiguity we had about the defocus value can be avoided by taking a through-focus series of micrographs. Phases that cannot be determined with certainty from one image may be evident in another one, taken at a different defocus. The deduced amplitudes and phases from several images may be compared and this will increase the certainty of the structure determination.

The high symmetry in this case was of course very helpful. It is, however, not absolutely necessary to take advantage of symmetry for solving a structure by CIP. In principle it should be possible to solve even a structure in the space group P1, by taking several pictures, with different optical parameters, especially a through-focus series. Every image will be the product of the structure plus a limited set of more or less unknown optical parameters. If the system is overdetermined, i.e. there are more reflections than unknown electron optical parameters for every picture, then the unknown parameters can in principle be determined and the crystal structure solved.

One very important factor in the solving of phases is the slow variation of phase errors in reciprocal space. This makes it possible to follow the errors as they propagate with increasing resolution. The larger the unit cell is, the closer together the lattice points will be in reciprocal space, and the easier it will be to follow the gradual

phase changes. A unit cell of dimensions above 2 nm like the present one is a great advantage in this procedure. Yet this is not absolutely necessary, but for considerably smaller unit cells a single picture will probably not suffice for doing this kind of analysis.

High valence metal oxides with open structures and one short axis are of course ideal objects for structure determination procedures by HREM, since the typical distances between the heavy metal atoms are 0.3–0.4 nm. Many niobium and tungsten oxides fulfill these requirements. In cases with shorter distances between metal atoms in projection, higher resolution is of course necessary. Similarly, in order to determine the positions of the lighter atoms, for example oxygen, higher resolution is needed. However, also in these cases CIP should be useful for correcting the electron optical distortions.

How useful can CIP be for investigations of non-periodic features, such as dislocations, faults or quasi-crystals? Clearly the method described here is not adequate for these cases. Yet after some modification CIP may well be very fruitful also in these cases, since the interpretations of non-periodic features are very dependent on good and, if possible, distortion-free images. In those cases where a non-periodic feature exists adjacent to a crystal on the same micrograph, the procedure can be applied to find the optical parameters, by analysing the crystalline area. The optical conditions are almost constant over the image so the optical parameters found for the crystalline region should be valid also in the other regions of the image. If the optical conditions are known the optical distortions can be corrected for, in reciprocal space, and not only at the lattice points, but continuously over the entire Fourier transform. The corrected Fourier transform can then be used to reconstruct the original scanned image, but now without the optical distortions. Such a picture would be very reliable in terms of contrast and relative positions of atoms.

13. Conclusions

By combining HREM and CIP the metal atom arrangement in thin metal oxides can be determined with high accuracy, even from only a single heavily distorted image, and without using a priori any chemical or structural knowledge. Atomic positions of heavy metals are correct to within about 10 pm if the resolution of the image is 0.25 nm. The different distortions caused by optical and crystal misalignments can be found and corrected for, giving an almost distortion-free density map of the crystal structure. If CIP is applied to images of even higher resolution, it should become possible to determine positions of atoms closer together, such as in alloys. With the modifications proposed here image processing should help in interpreting also non-crystalline features, such as dislocations and quasi-crystals.

Acknowledgements

The authors wish to thank Professor Peder Kierkegaard for support, Mikael Hovmöller for writing a computer program for comparing phases in P4bm and The Swedish Science Research Council (NFR) for financial support. The authors also gratefully acknowledge comments made by referee R. Henderson on the first version of the manuscript.

References

[1] J.D. DeRosier and A. Klug, Nature 217 (1968) 130.
[2] L.A. Amos, R. Henderson and P.N.T. Unwin, Progr. Biophys. Mol. Biol. 39 (1982) 183.
[3] S. Hovmöller, A. Sjögren, G. Farrants, M. Sundberg and B-O. Marinder, Nature 311 (1984) 238.
[4] D.J. Smith, W.O. Saxton, M.A. O'Keefe, G.J. Wood and W.M. Stobbs, Ultramicroscopy 11 (1983) 263.
[5] W.O. Saxton and D.J. Smith, Ultramicroscopy 18 (1985) 39.
[6] R. Henderson, J.M. Baldwin, K.H. Downing, J. Lepault and F. Zemlin, Ultramicroscopy 19 (1986) 147.
[7] B.M. Gatehouse, D.J. Lloyd and B.K. Miskin, Natl. Bur. Std. (US) Spec. Publ. 364 (1972) 15.
[8] M.A. O'Keefe, P.R. Buseck and S. Iijima, Nature 274 (1978) 322.
[9] J.M. Cowley and A.F. Moodie, Acta Cryst. 10 (1957) 609.
[10] V. Bhide and M. Gasperin, Acta Cryst. B35 (1979) 1318.
[11] M. Lundberg, Chem. Commun. Univ. Stockholm No. 12 (1971).
[12] D. Li and S. Hovmöller, J. Solid State Chem. 73 (1988) 5.