J. Appl. Cryst. (1988). **21**, 67–71

Automatic Indexing of Rotation Diffraction Patterns

BY WOLFGANG KABSCH

Max-Planck-Institut für Medizinische Forschung, Abteilung Biophysik, Jahnstrasse 29, D-6900 Heidelberg, Federal Republic of Germany

(*Received* 29 *April* 1987; *accepted* 1 *October* 1987)

Abstract

A method is described which assigns indices to a set of single-crystal reflections recorded by the rotation–oscillation technique using a fixed X-ray wavelength. It is assumed that the space group and approximate unit-cell parameters are known. The unknown crystal orientation is determined directly from the observed diffraction pattern of one or several oscillation data records. A local indexing procedure is described which tolerates large initial errors in the parameters controlling the diffraction pattern. These parameters are refined subsequently, thereby satisfying the constraints imposed by the space-group symmetry.

Introduction

During the course of macromolecular structure determination many crystals are required to collect native and heavy-atom derivative data. For each crystal the orientation of its unit cell must be determined. This may be a time-consuming process if the morphology of the crystal bears little relationship to its unit cell, as is often the case. In addition, a complicated alignment procedure may seriously affect the amount of useful data which can be collected if the crystal is sensitive to X-ray exposure. A method which automatically derives the unknown orientation of the crystal from the data records is thus desirable. The algorithm described in this paper was originally developed for processing data collected with a position-sensitive area detector but was found to be equally useful for conventional oscillation films.

Coordinate systems

1. *The orthonormal laboratory coordinate system*

This system $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is described in Fig. 1. The only alignment conditions are: (*a*) The detector-plane normal must be perpendicular to the spindle axis for all values of the swing angle χ. By definition, the vector $\mathbf{e}_3$ is parallel to the detector normal at $\chi = 0°$. (*b*) The direct beam must strike the crystal at a point lying on the spindle axis.

2. *Detector coordinate system*

A point X, Y on the detector plane is specified in raster steps with respect to two orthonormal vectors $\mathbf{e}_1^D, \mathbf{e}_2^D$ spanning the plane. At $\chi = 0°$, the projection of these vectors onto the xy plane in the laboratory system is specified by an angle ω. The origin is at X_0, Y_0 raster steps and marks the end point of the vector $\mathbf{F}$ (see Fig. 1). Q is the length of a raster step specified in the same units as the crystal–detector distance F. The relationships between the detector coordinates X, Y and the laboratory coordinates x, y, z of a point $\mathbf{x}$ on the detector plane are

$$\begin{aligned} x &= [(X - X_0)\cos\omega + (Y - Y_0)\sin\omega]Q\cos\chi \\ &\quad + F\sin\chi \\ y &= [(Y - Y_0)\cos\omega - (X - X_0)\sin\omega]Q \\ z &= -[(X - X_0)\cos\omega + (Y - Y_0)\sin\omega]Q\sin\chi \\ &\quad + F\cos\chi \end{aligned} \qquad (1)$$

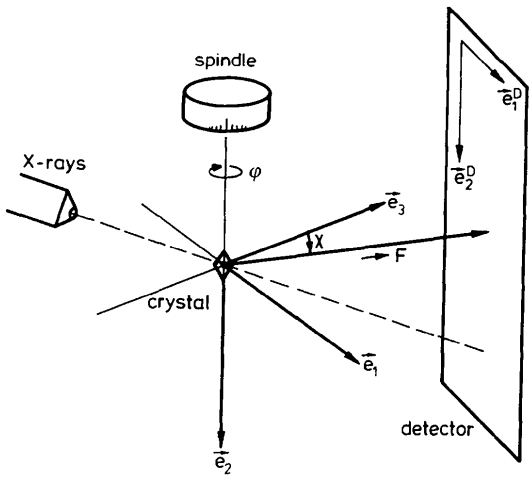


Fig. 1. Definition of the Cartesian laboratory coordinate system $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. The origin is on the intersection of the spindle axis with the primary X-ray beam which is also at the centre of the crystal. The detector can be rotated by an angle χ about an axis coinciding with the spindle. The $\mathbf{e}_3$ axis is collinear with the detector normal at swing angle $\chi = 0°$. Note that the primary beam is not required to be parallel to $\mathbf{e}_3$. The detector coordinate system is represented by the two orthonormal vectors $\mathbf{e}_1^D, \mathbf{e}_2^D$.

0021-8898/88/010067-05$03.00

and

$$X = X_0 + [(x\cos\chi - z\sin\chi)\cos\omega - y\sin\omega]/Q$$
$$Y = Y_0 + [(x\cos\chi - z\sin\chi)\sin\omega + y\cos\omega]/Q. \quad (2)$$

3. *Crystal coordinate system*

The reciprocal unit-cell vectors **a***, **b***, **c*** are represented with respect to the crystal coordinate system as columns of the matrix A_0 (Kabsch, 1977),

$$A_0 = \begin{pmatrix} a^*\sin\beta^* & b^*(\cos\gamma^* - \cos\alpha^*\cos\beta^*)/\sin\beta^* & 0 \\ 0 & b^*[\sin^2\alpha^* - (\cos\gamma^* - \cos\alpha^*\cos\beta^*)^2/\sin^2\beta^*]^{1/2} & 0 \\ a^*\cos\beta^* & b^*\cos\alpha^* & c^* \end{pmatrix}. \quad (3)$$

At spindle position $\varphi = 0°$ the representation of the matrix A_0 with respect to the laboratory frame is

$$A = UA_0, \quad (4)$$

where the orientation matrix U depends on three free parameters $\varepsilon_1, \varepsilon_2, \varepsilon_3$. These parameters are used to describe the effect of a small rotation on U as

$$U_{new} = \begin{pmatrix} 1 & -\varepsilon_3 & \varepsilon_2 \\ \varepsilon_3 & 1 & -\varepsilon_1 \\ -\varepsilon_2 & \varepsilon_1 & 1 \end{pmatrix} U_{old}. \quad (5)$$

If the crystal is rotated around the spindle by an angle φ the representation of A_0 with respect to the laboratory system changes to

$$D(\varphi)A = D(\varphi)UA_0 \quad (6)$$

where

$$D(\varphi) = \begin{pmatrix} \cos\varphi & 0 & \sin\varphi \\ 0 & 1 & 0 \\ -\sin\varphi & 0 & \cos\varphi \end{pmatrix}. \quad (7)$$

Diffraction geometry

The incident beam is represented by the wave vector

$$\mathbf{s}_0 = \begin{pmatrix} -(\sin\eta\cos\kappa)/\lambda \\ (\sin\kappa)/\lambda \\ (\cos\eta\cos\kappa)/\lambda \end{pmatrix} \quad (8)$$

with respect to the laboratory frame. λ denotes the fixed X-ray wavelength. The beam direction is described by the two angles η, κ. A reflection with indices h, k, l is represented as the vector

$$\mathbf{p} = \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = A\mathbf{h} \quad (9)$$

and

$$\mathbf{h} = \begin{pmatrix} h \\ k \\ l \end{pmatrix} \quad (10)$$

with A from (4). Rotation of the crystal around the spindle axis by φ may change **p** to a new vector **r** such that the Laue equation is satisfied:

$$\mathbf{r} = D(\varphi)\mathbf{p} = \mathbf{S} - \mathbf{S}_0. \quad (11)$$

The wave vector of the diffracted beam,

$$\mathbf{S} = \begin{pmatrix} S_x \\ S_y \\ S_z \end{pmatrix} \quad (12)$$

has the same magnitude as $\mathbf{S}_0$. The coordinates of **r** are

$$r_x = p_x\cos\varphi + p_z\sin\varphi = -\rho\sin(\tau - \eta)$$
$$r_y = p_y \quad (13)$$
$$r_z = p_z\cos\varphi - p_x\sin\varphi = -\rho\cos(\tau - \eta)$$

where

$$\rho = [p_x^2 + p_z^2]^{1/2} = [r_x^2 + r_z^2]^{1/2} \quad (14)$$
$$\cos\tau = (\mathbf{p}^2\lambda/2 + p_y\sin\kappa)/(\rho\cos\kappa) \quad (15)$$
$$\sin\tau = \pm[1 - \cos^2\tau]^{1/2}. \quad (16)$$

If the right-hand side of (15) is of magnitude less than 1 there are two solutions for $\sin\tau$, otherwise the reflection will be in the blind region. The angles φ at which diffraction occurs satisfy the conditions

$$\sin\varphi = (r_x p_z - r_z p_x)/\rho^2$$
$$\cos\varphi = (r_x p_x + r_z p_z)/\rho^2. \quad (17)$$

The diffracted beam hits the detector at

$$\mathbf{x} = \mathbf{S}F/(S_x\sin\chi + S_z\cos\chi), \quad (18)$$

the detector coordinates X, Y being determined by (2). Hence, the angular position φ and detector coordinates X, Y of a reflection h, k, l depend on the variables

$$X_0, Y_0, F, \kappa, \varepsilon_1, \varepsilon_2, \varepsilon_3, a^*, b^*, c^*, \alpha^*, \beta^*, \gamma^* \quad (19)$$

and constants

$$\lambda, \eta, \omega, \chi, Q. \quad (20)$$

Conversely, reflection indices can be calculated for a given X, Y, φ if the above parameters are known,

$$\mathbf{h} = [D(\varphi)UA_0]^{-1}(\mathbf{x}/|\lambda\mathbf{x}| - \mathbf{S}_0), \quad (21)$$

and **x** is obtained from (1).

Observed diffraction pattern

A list of n visible diffraction spots,

$$X_i, Y_i, \varphi_i, \quad i = 1, \ldots, n, \quad (22)$$

is obtained from oscillation data records by a pattern-recognition procedure developed earlier (Kabsch, 1977). The list (22) is sorted in order of increasing resolution by a non-recursive version of the *QUICKSORT* algorithm (Wirth, 1976). For each of the spots in the list (22) the coordinates of its corresponding diffracted-beam wave vector $\mathbf{S}_i$ are obtained from (1). The magnitude of $\mathbf{S}_i$ is set to $1/\lambda$. Solution of (11) thus yields a list of reciprocal-space vectors

$$\mathbf{p}_i = D(-\varphi_i)(\mathbf{S}_i - \mathbf{S}_0), \quad i = 1, \ldots, n \qquad (23)$$

corresponding to the observed spots. The unknown indices cannot be determined from (9) since the crystal orientation is not yet available.

Crystal orientation

The problem of finding the crystal orientation from a list of observed reciprocal-space lattice vectors such as (23) has been discussed by Sparks (1976) in the context of data collection with a diffractometer. Unfortunately none of the approaches can be used here directly because there may not be any very-low-resolution reflection of sufficient accuracy in the list (23). One should remember that this list may contain some artefacts since it is generated by an automatic routine and not carefully chosen by a crystallographer. However, as described below, a reliable list of very-low-resolution reciprocal-lattice points can be obtained by taking advantage of the large number of observed reflections on a rotation data record.

From the list (23) of observed reciprocal-lattice points low-resolution differences

$$\mathbf{p}_{ij} = \mathbf{p}_i - \mathbf{p}_j, \quad i < j, \quad i,j = 1, \ldots, n \qquad (24)$$

are formed and accumulated in a three-dimensional histogram. The grid size is set to one-tenth of the length of the shortest reciprocal-space cell axis. As expected, the low-resolution difference vectors nicely cluster in the histogram. The most densely occupied clusters are again located with the spot-finding procedure. The resulting list of cluster vectors is sorted into decreasing populations of accumulated difference vectors. The first two $\mathbf{v}_1$, $\mathbf{v}_2$ which are at an angular separation of larger than 45° are selected. The task of finding the orientation of the unit cell requires the correct assignment of indices to these two vectors. This is achieved by generating a list of very-low-order reciprocal-lattice vectors from the known unit-cell parameters. Only those vectors are omitted from the list which do not obey the general conditions limiting possible reflections. For example, in tetragonal space group $I4_1$ the reflection 002 would be kept but the reflection 001 would be discarded from the list since it can never occur as a difference vector between observed reflections. Indices $\mathbf{h}_1$, $\mathbf{h}_2$ from the generated list of possible reflections are assigned to the two observed vectors $\mathbf{v}_1, \mathbf{v}_2$ and a best rotation U is determined by minimizing (Kabsch, 1976)

$$E = (UA_0\mathbf{h}_1 - \mathbf{v}_1)^2 + (UA_0\mathbf{h}_2 - \mathbf{v}_2)^2. \qquad (25)$$

Since the space group is assumed to be known symmetry-related solutions are recognized and the six best independent assignments are kept. The remaining difference-vector clusters are indexed by the local indexing procedure described below.

Local indexing

Initially, some of the parameters controlling the diffraction pattern may have substantial errors. Refinement of these parameters requires a sufficiently large list of correctly indexed reflections and their observed positions on the detector.

Let $\mathbf{h}_i$ denote these yet unknown indices for each of the vectors $\mathbf{p}_i$ in the list of observations (23). Then the indices $\mathbf{h}_{ij}$ for each difference vector between reciprocal-lattice points i and j obey the relation

$$\mathbf{h}_{ij} = \mathbf{h}_i - \mathbf{h}_j. \qquad (26)$$

Because of the rotation–oscillation technique used for recording the diffraction pattern there are always many short difference vectors available. Their indices $\mathbf{h}_{ij}$ assume very small numbers which can be found using the matrix A_0 from (3) and an initial orientation matrix U as described in the last section. Then the unknown indices $\mathbf{h}_i$ can be found by minimizing the expression

$$H = \tfrac{1}{4}\sum_i \sum_j w_{ij}(\mathbf{h}_i - \mathbf{h}_j - \mathbf{h}_{ij})^2 + \tfrac{1}{2}\left(\sum_i \mathbf{h}_i\right)^2 \qquad (27)$$

where $w_{ij} = 1$ if the lattice points i and j are close, zero otherwise.

The second term was included to fix the origin of the indices $\mathbf{h}_i$. In fact the situation is slightly more complicated because the set of observed reciprocal-lattice points $\mathbf{p}_i$ may consist of several disconnected components. Then for each component an additional term is added to (27) to fix the arbitrary origin of its indices. Minimization of (27) leads to a well behaved linear system of equations, at least for dimensions up to $n = 350$ in single-precision (6–7 digits accuracy) arithmetic. The solution obtained consists of real-valued index vectors $\mathbf{h}_i$, their sum being zero for each connected set. The relative offsets between these sets of reciprocal-lattice points are found in the following way.

First, for each connected set μ its centroid $\mathbf{t}_\mu$ of reciprocal-lattice points $\mathbf{p}_i^\mu$ is determined and subtracted. A better orientation matrix U is then found by minimizing

$$\sum_\mu \sum_j [UA_0\mathbf{h}_j^\mu - (\mathbf{p}_j^\mu - \mathbf{t}_\mu)]^2. \qquad (28)$$

Table 1. *Results of automatic indexing of a single oscillation data film (space group and approximate unit-cell parameters were assumed to be known)*

Crystal: DNase I + self-complementary octanucleotide GCGATCGC; space group $C222_1$ $a = 72{\cdot}9$, $b = 100{\cdot}1$, $c = 92{\cdot}6$ Å.
Data: 1·5° oscillation film containing 1313 visible spots out to 1·9 Å resolution, digitized at 100 μm raster.
Results: 1213 spots have been indexed automatically: $\sigma_{XY} = 0{\cdot}114$ mm, $\sigma_{\varphi} = 0{\cdot}47°$.

Difference vector clusters						
Coordinates (Å^{-1})			Frequency	Logical indices		
x	y	z		h	k	l
−0·0003	0·0104	0·0000	278	−0·03	0·00	1·00
−0·0163	0·0001	0·0038	149	−1·00	1·00	0·00
−0·0160	−0·0106	0·0038	138	−1·00	1·00	−1·00
−0·0001	0·0212	0·0000	133	−0·03	0·00	2·00
−0·0161	0·0106	0·0038	120	−1·00	1·00	1·00
−0·0160	−0·0214	0·0038	95	−0·99	1·00	−2·00
0·0315	0·0001	0·0084	94	0·98	−3·00	0·00
0·0317	−0·0108	0·0084	91	0·98	−3·00	−1·00
0·0315	0·0107	0·0084	84	0·98	−3·00	1·00
−0·0164	0·0212	0·0038	82	−1·01	1·00	2·00

With the improved orientation matrix U the relative offsets

$$\mathbf{h}^{\mu} = (UA_0)^{-1}\mathbf{t}_{\mu} \tag{29}$$

are added to each $\mathbf{h}_j^{\mu}$. Finally, the offset common to all indices is determined and subtracted.

Refinement

Refinement of the parameters (19) is based upon the first hundred reflections in the sorted list (22) of observations indexed by the local indexing method. Better estimates of these parameters are then used to index spots at higher resolution according to a method described earlier (Milch & Minor, 1974; Kabsch, 1977). The number of newly indexed reflections is restricted to the number of spots used for refinement. The refinement is then repeated with the number of indexed observations doubling in each step.

Better estimates of the parameters are obtained by minimizing the expression

$$E = w_{XY}\sum_i[(\Delta_i^X)^2 + (\Delta_i^Y)^2] + w_{\varphi}\sum_i(\Delta_i^{\varphi})^2 \tag{30}$$

with spatial and angular weight constants w_{XY} and w_{φ}. Summations extend over all observed and indexed reflections. The residuals are

$$\begin{aligned}\Delta_i^X &= X_i - X_i^{\mathrm{obs}}\\ \Delta_i^Y &= Y_i - Y_i^{\mathrm{obs}}\\ \Delta_i^{\varphi} &= (\rho_i/|\mathbf{x}_i|)2\sin[(\varphi_i - \varphi_i^{\mathrm{obs}})/2]\end{aligned} \tag{31}$$

where X_i, Y_i, φ_i are calculated detector positions and spindle angles when reflection i is in the diffracting condition. The first factor in the expression for Δ_i^{φ} weighs down reflections close to the spindle axis. As in the method of least squares (*e.g.* Cruickshank, 1967) the residuals are expanded to first order in terms of the parameters (19). Minimization of (30) leads to a linear system of normal equations. Their solution provides a better estimate for the parameters (19). The derivatives of the residuals with respect to the independent parameters are obtained from the expressions given in this paper and are not presented here. Space-group symmetry may lead to additional relations between the unit-cell parameters. The derivatives of the matrix A_0 (3) with respect to the independent cell parameters are then obtained by use of the chain rule.

Results

The performance of the algorithm described in this paper was investigated under various conditions including simulated rotation data with random errors, area-detector data and conventional oscillation films.

Table 1 shows the results of applying the algorithm to a 1·5° oscillation data film of DNase I cocrystallized with the self-complementary octanucleotide GCGATCGC (Suck, Lahm & Oefner, 1988). The film was digitized at 100 μm raster and processed subsequently by the spot-finding procedure described earlier (Kabsch, 1977). The resulting list of 1313 diffraction-spot coordinates together with the unit-cell parameters, space-group number and crystal-to-film distance was presented to the automatic indexing procedure. For all spots their spindle positions at diffraction were assumed to be at the centre of the oscillation range of the film. A corresponding list of reciprocal-lattice vectors was calculated from (23). The ten most frequently occurring difference vectors are dis-

played in Table 1. Several plausible sets of indices were assigned to the first two of these and extended to the rest of the difference-vector clusters by the local indexing procedure. The best assignment is shown in Table 1. The small deviations of the indices from integral values indicate that some cluster vectors of lower population (not shown) are not fitting perfectly. Each of the possible solutions was refined. As a result a unique best solution emerged with a root-mean-square deviation of 0·114 mm in the spot locations and 0·47° in the spindle positions.

Automatic indexing was repeated using 80 Å instead of the correct value of 72·9 Å as an initial guess for the *a* axis. Despite this large error the observed spots were indexed correctly and the cell parameters refined to their correct values.

I am grateful to Dietrich Suck for his oscillation film and to Emil Pai and Ken Holmes for numerous discussions.

After submission of this paper it has come to my attention that A. Howard has developed a similar method for autoindexing based on short difference vectors (Howard 1986).

References

CRUICKSHANK, D. W. J. (1967). *International Tables for X-ray Crystallography*, edited by J. S. KASPER & K. LONSDALE, Vol. II, pp. 92–94. Birmingham: Kynoch Press. (Present distributor D. Reidel, Dordrecht.)

HOWARD, A. (1986). Proc. EEC Cooperative Workshop on Position-Sensitive Detector Software (Phases I & II), LURE, Paris, 26 May–7 June 1986, pp. 89–94.

KABSCH, W. (1976). *Acta Cryst.* A**32**, 922–923.

KABSCH, W. (1977). *J. Appl. Cryst.* **10**, 426–429.

MILCH, J. R. & MINOR, T. C. (1974). *J. Appl. Cryst.* **7**, 502–505.

SPARKS, R. A. (1976). In *Crystallographic Computing Techniques*, edited by F. R. AHMED, pp. 452–467. Copenhagen: Munksgaard.

SUCK, D., LAHM, A. & OEFNER, C. (1988). To be published.

WIRTH, N. (1976). *Algorithms + Data Structures = Programs*, pp. 80–82. New York: Prentice-Hall.

Short Communications

Contributions intended for publication under this heading should be expressly so marked; they should not exceed about 1000 *words; they should be forwarded in the usual way to the appropriate Co-editor; they will be published as speedily as possible.*

J. Appl. Cryst. (1988). **21**, 72–73

Evidence of a single-*q* incommensurate phase in quartz by synchrotron X-ray diffraction.

By A. ZARKA, B. CAPELLE and M. PETIT, *Laboratoire de Minéralogie–Cristallographie, Université Pierre et Marie Curie, Tour* 16, 4 *place Jussieu*, 75252 *Paris CEDEX* 05, *France*, and G. DOLINO, P. BASTIE and B. BERGE, *Laboratoire de Spectrométrie Physique, BP* 87, 38402 *Saint Martin d'Hères CEDEX, France*

(*Received* 27 *May* 1987; *accepted* 5 *August* 1987)

Abstract

X-ray scattering is used to demonstrate the existence in quartz of an incommensurate phase with a single modulation when a uniaxial stress is applied in the XY plane. Good agreement with earlier neutron scattering experiments is found.

Introduction

Experimental evidence for an incommensurate phase in quartz between the α and β phases has been given by Berge, Dolino, Vallade, Boissier & Vacher (1984); Dolino, Bachheimer, Berge & Zeyen (1984); Gouhara, Li & Kato (1983*a*, *b*); Gouhara & Kato (1985*a*, *b*); Van Tendeloo, Van Landuyt & Amelinckx (1976) and Van Tendeloo (1987). A Landau-type theory gives two possibilities for its structure: in the structure called 3*q*, three modulations at 120° from each other in the XY plane are superimposed everywhere in the crystal; in the one called 1*q*, they are spatially separated, and only one modulation exists in each point. Only the 3*q* structure has been observed. The same theory predicts that a 3*q*–1*q* transition occurs when a uniaxial stress is applied in the XY plane. This transition has actually been observed by neutron scattering (Dolino, Bastie, Berge, Vallade, Bethke, Regnault & Zeyen, 1987). In this paper we describe X-ray scattering experiments that confirm the occurrence of this transition.

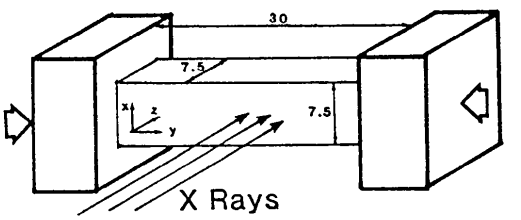
Experimental

All experiments were carried out using synchrotron radiation from LURE in Orsay and a furnace working at atmospheric pressure. We used natural quartz crystals. We performed two sets of experiments. In the first set, using a device developed for neutron scattering we applied a compressive stress along the Y axis and we used a reflection Laue setup (Fig. 1*a*). For the second set, we built a device designed to apply traction. This device allows thinner plates, so that we could use a transmission Laue setup (Fig. 1*b*) (Gouhara *et al.* 1983*a*).

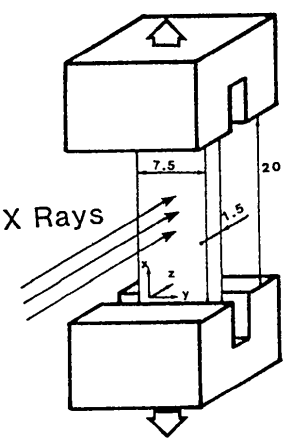
Results and discussion

The device and the setting of the crystal we used in the first set of experiments allow applied compressive stresses up to 60 MPa. The photographs were taken around the $01\bar{3}$ Bragg spot. They were obtained in 15–20 min on Kodak dental film.

In good agreement with theory and previous neutron scattering experiments, we found a linear dependence of the α-incommensurate and β-incommensurate transition temperatures on applied stress. At constant applied stress, as the



(*a*)



(*b*)

Fig. 1. Setting of the crystal and direction of the applied stress. Dimensions are given in mm. (*a*) First set of experiments; (*b*) second set of experiments.

0021-8898/88/010072-02$03.00