Acta Cryst. (1990). A**46**, 425–435

On the Use of Normal Modes in Thermal Parameter Refinement: Theory and Application to the Bovine Pancreatic Trypsin Inhibitor

BY R. DIAMOND

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England

(*Received* 3 *October* 1989; *accepted* 5 *February* 1990)

Abstract

A method is presented whereby the amplitude coefficients of molecular normal modes of vibration are treated as independent variables in the treatment of thermal effects in X-ray diffraction, and applied to the bovine pancreatic trypsin inhibitor, form II ($P2_12_12_1$, a=74·1, b=23·4, c=28·9 Å). It is shown that the description of molecular motion furnished by 892 isotropic temperature factors may be largely reproduced using only 19 molecular thermal parameters from which anisotropic temperature factors may be synthesised for every atom. The method shows that motions and/or disorders external to each molecule are the largest single source of apparent motion, and that the internal motions are comparable to those predicted by Levitt, Sander & Stern [*J. Mol. Biol.* (1985). **181**, 423–427] at 300 K.

1. Introduction

It has become increasingly evident in recent years that the conventional treatment of thermal motion in macromolecules using isotropic Debye–Waller factors provides only a poor description of the thermal motions believed to be present in such molecules, and the use of anisotropic Debye–Waller factors for proteins is usually precluded by the adverse ratio of parameters to observations that results when this is attempted. Some proteins, including the bovine pancreatic trypsin inhibitor (BPTI) (Wlodawer, Walter, Huber & Sjölin, 1984), have, however, been refined with anisotropic thermal parameters, using the method of Konnert & Hendrickson (1980) and Hendrickson & Konnert (1981) which imposes restraints that ensure that the thermal parameters of neighbouring atoms be similar, thus overcoming, to some extent, the shortage of data. Howlin, Moss & Harris (1989) have also performed a refinement of anisotropic thermal parameters for selected rigid-body groupings of bovine ribonuclease-A using the **TLS** method of Schomaker & Trueblood (1968). Both these methods give results which are more detailed than the present work since their models possess more parameters. The objective here, however, has been to test whether the salient features of observed B factors are accountable in terms of a combination of internal modes of vibration and rigid-body motions of the entire molecule and to attempt an apportionment of the apparent motions to each of these sources.

Attempts to describe macromolecular thermal motion in a more satisfactory manner have arisen in the field of molecular mechanics in which there are two main approaches. One of these consists of integrating the equations of motion of each and every atom in the molecule using empirical interatomic force fields. The immediate result of such a calculation is a very large number of coordinate sets and velocity sets separated one from the next by some small interval of time. From the statistics of such sets it is possible to extract mean atomic positions and moments of the probability density function (p.d.f.) of atomic centres about such means. Correlations between motions of neighbouring atoms may also be studied. The strength of this method is that it makes no assumptions concerning the nature of the atomic motions and it is capable of modelling situations in which atoms move in anharmonic potential wells and have p.d.f.s which are far from Gaussian. The limitations on its applicability arise from the scale of the calculations involved and from the fact that it is difficult to close any feedback loop to improve the potential functions that govern the simulation in the event that observed and calculated diffraction amplitudes fail to agree.

The other main approach consists of the estimation of the normal modes of vibration of the molecule from the matrix of second derivatives of potential energy at the potential-energy minimum. This approach is complementary to the first. By working with internal coordinates, especially dihedral angles, the dominant motions, *i.e.* those with low frequency and high amplitude, may be characterized using many fewer modes than the $3n-6$ for n atoms which are required to describe the elastic character of the molecule fully (Levitt, Sander & Stern, 1985). The main limitations of the approach are that it is limited to the harmonic approximation, and that it cannot properly deal with bound water molecules. Nevertheless it is assumed in this work that the motions actually present in the molecule, and which are simulated by the molecular-dynamics approach, contain strong components of motions attributable to the low-frequency high-amplitude modes, and the method described below estimates amplitude coefficients for a given set of modes directly from the X-ray observations. When these coefficients have been

0108-7673/90/060425-11$03.00

determined, anisotropic Debye–Waller factors for every atom may be synthesized from them, giving a description of the thermal motion based on a small number of independent variables, which may then be compared with more traditional treatments.

2. Theory

We begin by reviewing the theory so that the assumptions implicit in the treatment are exhibited. The instantaneous scattering amplitude for the crystal is

$$F(\mathbf{s}) = \sum_i f_i \exp[2\pi i(\mathbf{q}_i + \mathbf{r}_i).\mathbf{s}] \tag{1}$$

for scattering vector $\mathbf{s}$, [$|\mathbf{s}| = (2\sin\theta)/\lambda$] where $\mathbf{r}_i$ is the fixed mean position of atom i and $\mathbf{q}_i$ is its instantaneous displacement. The intensity is then given by

$$\begin{aligned} I(\mathbf{s}) &= \langle F(\mathbf{s})F^*(\mathbf{s})\rangle \\ &= \sum_i \sum_j f_i f_j \exp[2\pi i(\mathbf{r}_i - \mathbf{r}_j).\mathbf{s}] \\ &\quad \times \langle \cos 2\pi(\mathbf{q}_i - \mathbf{q}_j).\mathbf{s} + i \sin 2\pi(\mathbf{q}_i - \mathbf{q}_j).\mathbf{s}\rangle \end{aligned} \tag{2}$$

in which $\langle\rangle$ denotes time averaging.

The distribution of $(\mathbf{q}_i - \mathbf{q}_j)$ is assumed to be symmetrical so that the time average of the sine vanishes. Treating all vectors as column vectors and using a superscript T to denote transposition, we may define the six-dimensional vector

$$\mathbf{q}^T = (\mathbf{q}_i^T, \mathbf{q}_j^T) \tag{3}$$

for one ij pair; then

$$\begin{aligned} &\langle \cos 2\pi(\mathbf{q}_i - \mathbf{q}_j)^T\mathbf{s}\rangle \\ &\quad = |\mathbf{M}|^{-1/2} \int\int \exp(-\pi\mathbf{q}^T\mathbf{M}^{-1}\mathbf{q}) \\ &\qquad \times \cos 2\pi(\mathbf{q}_i - \mathbf{q}_j)^T\mathbf{s}\, dv_i dv_j \end{aligned} \tag{4}$$

in which the 6×6 matrix $\mathbf{M}$ describes the probability density in $\mathbf{q}$ space that atoms i and j are simultaneously displaced by $\mathbf{q}_i$ and $\mathbf{q}_j$ respectively, and $|\mathbf{M}|$ is its determinant. $\mathbf{M}$ is related to the fluctuations in $\mathbf{q}$ by

$$\begin{aligned} &|\mathbf{M}|^{-1/2}\exp(-\pi\mathbf{q}^T\mathbf{M}^{-1}\mathbf{q}) \\ &\quad = (2\pi)^{-3}|\langle\mathbf{q}\mathbf{q}^T\rangle|^{-1/2}\exp(-\tfrac{1}{2}\mathbf{q}^T\langle\mathbf{q}\mathbf{q}^T\rangle^{-1}\mathbf{q}) \end{aligned} \tag{5}$$

since $\langle\mathbf{q}\mathbf{q}^T\rangle$ contains variances and covariances of the elements of $\mathbf{q}$, so that

$$\mathbf{M} = 2\pi\langle\mathbf{q}\mathbf{q}^T\rangle. \tag{6}$$

Definition of $\boldsymbol{\Delta}$ as the 6×3 matrix

$$\boldsymbol{\Delta} = \begin{pmatrix} \mathbf{I} \\ -\mathbf{I} \end{pmatrix} \tag{7}$$

where $\mathbf{I}$ is the 3×3 identity gives

$$(\mathbf{q}_i - \mathbf{q}_j)^T = \mathbf{q}^T\boldsymbol{\Delta} \tag{8}$$

and

$$\begin{aligned} &\langle \cos 2\pi(\mathbf{q}_i - \mathbf{q}_j)^T\mathbf{s}\rangle \\ &\quad = |\mathbf{M}|^{-1/2}\int \exp(-\pi\mathbf{q}^T\mathbf{M}^{-1}\mathbf{q})\cos 2\pi\mathbf{q}^T\boldsymbol{\Delta}\mathbf{s}\, dv_{\mathbf{q}} \\ &\quad = \exp(-\pi\mathbf{s}^T\boldsymbol{\Delta}^T\mathbf{M}\boldsymbol{\Delta}\mathbf{s}). \end{aligned} \tag{9}$$

The model we adopt to describe the atomic motions within the crystal consists of two parts. The first part regards each molecule as having internal motions described by the normal modes calculated by Levitt, Sander & Stern (1985). Each such mode has a time dependence which is uniform within one molecule, but we suppose that the phases of equivalent modes in separate molecules are unrelated although their amplitudes and frequencies are treated as being the same. The second part consists of rigid-body displacements and rotations of each molecule in a manner which can be described in terms of long-period lattice modes which span the entire crystal.

If there are n internal modes active then we may write

$$\mathbf{q}_i = (\mathbf{N}_i \quad \mathbf{L}_i)\begin{pmatrix} \mathbf{t}_i \\ \mathbf{T}_i \end{pmatrix} \qquad \mathbf{q}_j = (\mathbf{N}_j \quad \mathbf{L}_j)\begin{pmatrix} \mathbf{t}_j \\ \mathbf{T}_j \end{pmatrix} \tag{10}$$

in which $\mathbf{N}_i$ and $\mathbf{N}_j$ have three rows each and n columns which contain the vector displacements for atoms i and j for unit amplitude of each mode. These matrices are considered to be a molecular property and not to vary from molecule to molecule. The $\mathbf{t}$ matrices are column vectors given by

$$\begin{aligned} \mathbf{t}^T = \big(&A_1\cos(\omega_1 t + \varphi_1),\ A_2\cos(\omega_2 t + \varphi_2), \ldots, \\ &A_n\cos(\omega_n t + \varphi_n)\big) \end{aligned} \tag{11}$$

which define the amplitude and time dependence of each modal displacement and include a phase angle φ for each mode. Each molecule is considered to have its own set of phases φ, and solely for this reason $\mathbf{t}_i$ is distinguished from $\mathbf{t}_j$, the distinction vanishing if atoms i and j are in the same molecule.

The matrices $\mathbf{L}_i$ and $\mathbf{L}_j$ are similar to $\mathbf{N}$ but describe the overall displacements of the molecules containing atoms i and j and have similar time-dependent vectors $\mathbf{T}_i$ and $\mathbf{T}_j$. For an ideally resonant system $\mathbf{T}_i = \mathbf{T}_j$ even if atoms i and j are widely separated (many unit cells apart) and the modes described by $\mathbf{L}_i$ and $\mathbf{L}_j$ span the entire crystal. However no such lattice modes are available.

Using (3) and (10) we evaluate $\langle \mathbf{q}\mathbf{q}^T \rangle$ as

$$\langle \mathbf{q}\mathbf{q}^T \rangle = \begin{pmatrix} (\mathbf{N}_i \quad \mathbf{L}_i) \begin{pmatrix} \langle \mathbf{t}_i\mathbf{t}_i^T \rangle & \langle \mathbf{t}_i\mathbf{T}_i^T \rangle \\ \langle \mathbf{T}_i\mathbf{t}_i^T \rangle & \langle \mathbf{T}_i\mathbf{T}_i^T \rangle \end{pmatrix} \begin{pmatrix} \mathbf{N}_i^T \\ \mathbf{L}_i^T \end{pmatrix} & (\mathbf{N}_i \quad \mathbf{L}_i) \begin{pmatrix} \langle \mathbf{t}_i\mathbf{t}_j^T \rangle & \langle \mathbf{t}_i\mathbf{T}_j^T \rangle \\ \langle \mathbf{T}_i\mathbf{t}_j^T \rangle & \langle \mathbf{T}_i\mathbf{T}_j^T \rangle \end{pmatrix} \begin{pmatrix} \mathbf{N}_j^T \\ \mathbf{L}_j^T \end{pmatrix} \\ (\mathbf{N}_j \quad \mathbf{L}_j) \begin{pmatrix} \langle \mathbf{t}_j\mathbf{t}_i^T \rangle & \langle \mathbf{t}_j\mathbf{T}_i^T \rangle \\ \langle \mathbf{T}_j\mathbf{t}_i^T \rangle & \langle \mathbf{T}_j\mathbf{T}_i^T \rangle \end{pmatrix} \begin{pmatrix} \mathbf{N}_i^T \\ \mathbf{L}_i^T \end{pmatrix} & (\mathbf{N}_j \quad \mathbf{L}_j) \begin{pmatrix} \langle \mathbf{t}_j\mathbf{t}_j^T \rangle & \langle \mathbf{t}_j\mathbf{T}_j^T \rangle \\ \langle \mathbf{T}_j\mathbf{t}_j^T \rangle & \langle \mathbf{T}_j\mathbf{T}_j^T \rangle \end{pmatrix} \begin{pmatrix} \mathbf{N}_j^T \\ \mathbf{L}_j^T \end{pmatrix} \end{pmatrix} \tag{12}$$

in which

$$\begin{aligned} \langle \mathbf{t}_i\mathbf{t}_i^T \rangle &= \langle \mathbf{t}_j\mathbf{t}_j^T \rangle \\ &= \mathbf{P} = \mathrm{diag}(A_1^2/2,\ A_2^2/2,\ \ldots\ A_n^2/2) \\ \langle \mathbf{T}_i\mathbf{T}_i^T \rangle &= \langle \mathbf{T}_j\mathbf{T}_j^T \rangle = \mathbf{Q} \\ &= \mathrm{diag}(A_{n+1}^2/2,\ A_{n+2}^2/2,\ \ldots\ A_{n+m}^2/2) \\ \langle \mathbf{T}_i\mathbf{T}_j^T \rangle &= \mathbf{R} \\ \langle \mathbf{T}_j\mathbf{t}_j^T \rangle &= \langle \mathbf{T}_i\mathbf{t}_j^T \rangle = \langle \mathbf{T}_i\mathbf{t}_i^T \rangle = \mathbf{O} \\ \langle \mathbf{t}_i\mathbf{t}_j^T \rangle &= \mathbf{P}\mathbf{C} \\ \mathbf{C} &= \mathrm{diag}(\cos\Delta\varphi_1,\ \cos\Delta\varphi_2,\ \ldots\cos\Delta\varphi_n) \end{aligned} \tag{13}$$

so that (9) becomes, from (6),

$$\begin{aligned} &\langle \cos 2\pi(\mathbf{q}_i - \mathbf{q}_j)^T\mathbf{s} \rangle \\ &= \exp[-2\pi^2\mathbf{s}^T(\mathbf{N}_i\mathbf{P}\mathbf{N}_i^T + \mathbf{L}_i\mathbf{Q}\mathbf{L}_i^T - \mathbf{N}_i\mathbf{P}\mathbf{C}\mathbf{N}_j^T \\ &\quad - \mathbf{L}_i\mathbf{R}\mathbf{L}_j^T - \mathbf{N}_j\mathbf{P}\mathbf{C}\mathbf{N}_i^T - \mathbf{L}_j\mathbf{R}^T\mathbf{L}_i^T \\ &\quad + \mathbf{N}_j\mathbf{P}\mathbf{N}_j^T + \mathbf{L}_j\mathbf{Q}\mathbf{L}_j^T)\mathbf{s}] \end{aligned} \tag{14}$$

which may then be inserted in (2) to give the intensity.

In principle, the expression for the intensity is a double summation and must be treated as such unless terms in (14) involving both i and j can be ignored. There are four such terms, two involving **C** and two involving **R**. If atoms i and j are in the same molecule **C** is the identity and the four terms involving **P** are $(\mathbf{N}_i - \mathbf{N}_j)\mathbf{P}(\mathbf{N}_i^T - \mathbf{N}_j^T)$. However, the double summation must be taken over the whole crystal (or the region of it in which the X-irradiation has phase coherence) so that the vast majority of ij pairs involve atoms in different molecules for which the expectation value of **C** is zero. If there are N molecules within the region of X-ray phase coherence, and if there is no phase coherence between corresponding molecular modes in different molecules then the summation in (2) may be written as $A + B$ where A is the summation over all atom pairs in which both members of the pair are in the same molecule, and B is the summation in which atoms i and j are in different molecules. A is N times the double summation over a single representative molecule with $\mathbf{C} = \mathbf{I}$, and B is $N(N-1)$ times the double summation over a single representative molecule with $\mathbf{C} = \mathbf{O}$. Terms of the first type are therefore present in proportion $(N-1)^{-1}$ and may be ignored. Some degree of local phase coherence among neighbouring molecules cannot alter this conclusion unless their number becomes comparable with N.

Similarly the matrix **R** is equal to **Q** if the crystal is treated as one elastic resonator and **L** contains a sufficiently large number of modes, but here the motions of unit cells remote from each other are treated as uncorrelated, **L** being used to describe the positional and orientational uncertainty of widely separated unit cells arising from any cause, including static disorder, and under this interpretation the expectation of **R** is also zero.

The eight terms in (14) are thus reduced to four and (2) becomes

$$\begin{aligned} I(\mathbf{s}) &= \sum_i \sum_j f_i f_j \exp[2\pi i(\mathbf{r}_i - \mathbf{r}_j)^T\mathbf{s}] \\ &\quad \times \exp[-2\pi^2\mathbf{s}^T(\mathbf{N}_i\mathbf{P}\mathbf{N}_i^T + \mathbf{L}_i\mathbf{Q}\mathbf{L}_i^T \\ &\quad + \mathbf{N}_j\mathbf{P}\mathbf{N}_j^T + \mathbf{L}_j\mathbf{Q}\mathbf{L}_j^T)\mathbf{s}] \\ &= \Big| \sum_j f_j \exp 2\pi i\mathbf{r}_j.\mathbf{s} \\ &\quad \times \exp[-2\pi^2\mathbf{s}^T(\mathbf{N}_j\mathbf{P}\mathbf{N}_j^T + \mathbf{L}_j\mathbf{Q}\mathbf{L}_j^T)\mathbf{s}] \Big|^2 . \end{aligned} \tag{15}$$

This reduction to the square of a single sum is possible only if **C** and **R** are treated as zero. With this reduction, the distinction between **N** and **L** and between **P** and **Q** need no longer be retained, $\mathbf{N}_j$ from now on being given $(n+m)$ columns for the n internal modes and the m lattice or external modes. The symbol **P** will henceforth stand for diag(**P**, **Q**), and elements on its diagonal will be given the symbol p (in place of $A^2/2$) and subscripted p or q.

Thus atoms contribute to each structure factor with temperature factor $\exp(-2\pi^2\mathbf{s}^T\mathbf{N}_j\mathbf{P}\mathbf{N}_j^T\mathbf{s})$ and $F(\mathbf{s})$ transforms to an image of stationary atoms each convoluted with the transform of this. Thus if the stationary-atom scattering factors are represented by expressions of the form

$$f_j = \sum_{k=1}^{K} Z_{jk} \exp(-\pi a_{jk}^2|\mathbf{s}|^2) \tag{16}$$

for which a suitable value of K might be 2 or 3, then the electron density is given by

$$\begin{aligned} \rho(\mathbf{r}) &= \sum_j \sum_k Z_{jk}|\mathbf{M}_{jk}|^{-1/2} \\ &\quad \times \exp[-\pi(\mathbf{r}^T - \mathbf{r}_j^T)\mathbf{M}_{jk}^{-1}(\mathbf{r} - \mathbf{r}_j)] \end{aligned} \tag{17}$$

in which

$$\mathbf{M}_{jk} = a_{jk}^2\mathbf{I} + 2\pi\mathbf{N}_j\mathbf{P}\mathbf{N}_j^T \tag{18}$$

where $\mathbf{M}_{jk}$ is now one of K 3×3 matrices associated with atom j, being a partition of $\mathbf{M}$, augmented on the diagonal; $\mathbf{r}_j$ is the position of atom j and Z_{jk} and a_{jk} are respectively the electron count and radius of the kth component of atom j so that the matrices $\mathbf{M}_{jk}$ incorporate a stationary-atom form factor.

To optimize the matrix $\mathbf{P}$ we require F_c and the derivatives $\partial F_c/\partial p$ for all reflections and all modes. These are obtained by constructing maps of $\rho(\mathbf{r})$ and of the function $\partial\rho(\mathbf{r})/\partial p$, one for each mode, and transforming these to reciprocal space. A full-matrix least-squares solution for $\mathbf{P}$ is then obtained in which the normal matrix is constructed numerically from these derivatives. An alternative analytical expression for the normal matrix elements, obtained by replacing summation over all reflections by analytic integrals over reciprocal space, has also been obtained but has not been used because double summation over close atom pairs is still required. Similarly, direct expressions for $\partial F_c/\partial p$ are readily obtained but appear computationally less efficient than the route indicated.

From (17)

$$\begin{aligned}\frac{\partial\rho(\mathbf{r})}{\partial p_p} = \sum_j\sum_k Z_{jk}\Big[&-\frac{1}{2}|\mathbf{M}_{jk}|^{-3/2}\frac{\partial|\mathbf{M}_{jk}|}{\partial p_p} \\ &+ |\mathbf{M}_{jk}|^{-1/2}\frac{\partial[-\pi(\mathbf{r}^T-\mathbf{r}_j^T)\mathbf{M}_{jk}^{-1}(\mathbf{r}-\mathbf{r}_j)]}{\partial p_p}\Big] \\ &\times\exp[-\pi(\mathbf{r}^T-\mathbf{r}_j^T)\mathbf{M}_{jk}^{-1}(\mathbf{r}-\mathbf{r}_j)].\end{aligned} \tag{19}$$

But

$$\partial\mathbf{M}_{jk}^{-1}/\partial p_p = -\mathbf{M}_{jk}^{-1}\mathbf{n}_{jp}2\pi\mathbf{n}_{jp}^T\mathbf{M}_{jk}^{-1} \tag{20}$$

in which $\mathbf{n}_{jp}$ is the pth column of $\mathbf{N}_j$, and

$$|\mathbf{M}_{jk}|^{-1}\partial|\mathbf{M}_{jk}|/\partial p_p = 2\pi\mathbf{n}_{jp}^T\mathbf{M}_{jk}^{-1}\mathbf{n}_{jp} \tag{21}$$

so that

$$\begin{aligned}&\partial\rho(\mathbf{r})/\partial p_p \\ &= \sum_j\sum_k Z_{jk}\Big[-\tfrac{1}{2}.2\pi\mathbf{n}_{jp}^T\mathbf{M}_{jk}^{-1}\mathbf{n}_{jp} \\ &\quad + \pi(\mathbf{r}^T-\mathbf{r}_j^T)\mathbf{M}_{jk}^{-1}\mathbf{n}_{jp}2\pi\mathbf{n}_{jp}^T \\ &\quad \times\mathbf{M}_{jk}^{-1}(\mathbf{r}-\mathbf{r}_j)\Big]|\mathbf{M}_{jk}|^{-1/2} \\ &\quad \times\exp[-\pi(\mathbf{r}^T-\mathbf{r}_j^T)\mathbf{M}_{jk}^{-1}(\mathbf{r}-\mathbf{r}_j)] \\ &= \pi\sum_j\sum_k Z_{jk}\Big\{2\pi[(\mathbf{r}^T-\mathbf{r}_j^T)\mathbf{M}_{jk}^{-1}\mathbf{n}_{jp}]^2 \\ &\quad - \mathbf{n}_{jp}^T\mathbf{M}_{jk}^{-1}\mathbf{n}_{jp}\Big\}|\mathbf{M}_{jk}|^{-1/2} \\ &\quad \times\exp[-\pi(\mathbf{r}^T-\mathbf{r}_j^T)\mathbf{M}_{jk}^{-1}(\mathbf{r}-\mathbf{r}_j)].\end{aligned} \tag{22}$$

3. Implementation

The parameters which are optimized are the diagonal elements of $\mathbf{P}$, (13), and an overall scale factor, and the least-squares operation is structured in a manner which permits phases φ [unrelated to those of (11)] associated with F_{obs} to be used if these are known and which recognizes that, on the complex plane, F_{obs} may be well defined in the radial direction, and less well defined in the tangential direction. This structuring provides for a number of techniques with the same code.

For each reflection we treat F as a two-dimensional real vector, rather than as a complex number, and write

$$\mathbf{F}_o(\mathbf{s}) = \begin{pmatrix}|F_o|\cos\varphi\\|F_o|\sin\varphi\end{pmatrix} \qquad \mathbf{F}_c(\mathbf{s}) = \begin{pmatrix}|F_c|\cos\alpha\\|F_c|\sin\alpha\end{pmatrix}$$

$$\frac{\partial\mathbf{F}_c(\mathbf{s})}{\partial p_p} = \begin{pmatrix}|G_p|\cos\beta_p\\|G_p|\sin\beta_p\end{pmatrix}. \tag{23}$$

Then, for scale factor S, the residual vector (one reflection) after adding parametric increments δS and $\boldsymbol{\delta}\mathbf{p}$ is

$$\boldsymbol{\delta} = \mathbf{d} - \mathbf{D}\begin{pmatrix}\delta S\\\boldsymbol{\delta}\mathbf{p}\end{pmatrix} \tag{24}$$

in which

$$\mathbf{d} = (\mathbf{F}_o - S\mathbf{F}_c)$$

$$\mathbf{D} = \begin{pmatrix}|F_c|\cos\alpha & S|G_1|\cos\beta_1 & \dots & S|G_{m+n}|\cos\beta_{m+n}\\|F_c|\sin\alpha & S|G_1|\sin\beta_1 & \dots & S|G_{m+n}|\sin\beta_{m+n}\end{pmatrix}. \tag{25}$$

The residual to be minimized is

$$E = \sum_{\mathbf{s}}\boldsymbol{\delta}^T\mathbf{W}\boldsymbol{\delta} \tag{26}$$

in which the weight matrix

$$\begin{aligned}\mathbf{W} &= \begin{pmatrix}\cos\varphi & -\sin\varphi\\\sin\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix}\sigma_r^{-2} & 0\\0 & \sigma_t^{-2}\end{pmatrix}\begin{pmatrix}\cos\varphi & \sin\varphi\\-\sin\varphi & \cos\varphi\end{pmatrix} \\ &= \begin{pmatrix}\sigma_r^{-2}\cos^2\varphi+\sigma_t^{-2}\sin^2\varphi & (\sigma_r^{-2}-\sigma_t^{-2})\sin\varphi\cos\varphi\\(\sigma_r^{-2}-\sigma_t^{-2})\sin\varphi\cos\varphi & \sigma_r^{-2}\sin^2\varphi+\sigma_t^{-2}\cos^2\varphi\end{pmatrix}\end{aligned} \tag{27}$$

is used to express the fact that $\mathbf{F}_o$ has variances σ_r^2 and σ_t^2 in the radial and tangential directions respectively (Diamond 1976). E is then minimized by

$$\begin{pmatrix}\delta S\\\boldsymbol{\delta}\mathbf{p}\end{pmatrix} = \Big(\sum_{\mathbf{s}}\mathbf{D}^T\mathbf{W}\mathbf{D}\Big)^{-1}\Big(\sum_{\mathbf{s}}\mathbf{D}^T\mathbf{W}\mathbf{d}\Big) \tag{28}$$

Table 1. *Coefficients a_{jk} of equation* (16)

Atom	H	C	N	O	P	S
Z_{j1}	0·60738	3·47518	3·94769	4·08719	8·86185	6·86537
Z_{j2}	0·38981	2·50334	3·03812	3·90173	3·50029	9·12697
Z_{j3}					2·63067	
a_{j1} (Å)	1·54017	1·45475	1·25908	1·16993	0·30335	1·43218
a_{j2} (Å)	0·74435	0·35283	0·38481	0·41408	1·91711	0·28788
a_{j3} (Å)					1·17225	
R.m.s. error	0·00249	0·02004	0·01266	0·00886	0·00588	0·00664

so that a typical element of the normal matrix is

$$S^2 \sum_{\mathbf{s}} |G_p||G_q| \left[\sigma_r^{-2} \cos(\beta_p - \varphi)\cos(\beta_q - \varphi) + \sigma_t^{-2} \sin(\beta_p - \varphi)\sin(\beta_q - \varphi)\right] \tag{29}$$

and of the gradient vector is

$$S \sum_{\mathbf{s}} |G_p| \{\sigma_r^{-2} \cos(\beta_p - \varphi)\left[|F_o| - S|F_c| \cos(\alpha - \varphi)\right] - \sigma_t^{-2} S|F_c| \sin(\beta_p - \varphi)\sin(\alpha - \varphi)\}. \tag{30}$$

Equation (28) is evaluated as

$$\begin{pmatrix} \delta S \\ \delta \mathbf{p} \end{pmatrix} = \mathbf{U}^{-1}\mathbf{A}\mathbf{Z}\mathbf{\Lambda}^{-1}\mathbf{A}^T\mathbf{U}^{T-1}\left(\sum_{\mathbf{s}} \mathbf{D}^T\mathbf{W}\mathbf{d}\right) \tag{31}$$

in which **U** is diagonal and such that

$$\mathbf{U}^{T-1}\left(\sum_{\mathbf{s}} \mathbf{D}^T\mathbf{W}\mathbf{D}\right)\mathbf{U}^{-1} \tag{32}$$

has 1 at every diagonal position, *i.e.* is the correlation matrix, and **A** is orthogonal and such that

$$\mathbf{\Lambda} = \mathbf{A}^T\mathbf{U}^{T-1}\left(\sum_{\mathbf{s}} \mathbf{D}^T\mathbf{W}\mathbf{D}\right)\mathbf{U}^{-1}\mathbf{A} \tag{33}$$

is diagonal and contains the eigenvalues of the correlation matrix. **Z** is a diagonal filter matrix, normally containing 1s but containing elements in the range 0 to 1 in positions corresponding to very small eigenvalues.

The purpose of **U** is to form a normal matrix which is independent of the scales on which the modal displacements in **N** and **L** are presented and which reflects the observational correlation among the modes. Off-diagonal elements in $\mathbf{U}^{T-1}\left(\sum_{\mathbf{s}} \mathbf{D}^T\mathbf{W}\mathbf{D}\right)\mathbf{U}^{-1}$ frequently reach 0·6 and are all positive except in the first row and column where they are all negative, in accordance with the well known interaction of scale and temperature factors.

Diagonalization by **A** has been included because these high correlations lead to a wide eigenvalue spectrum, even after scaling by **U**. So far, however, condition numbers ($\lambda_{max}/\lambda_{min}$) have been about 50 and it has not been necessary to invoke this mechanism, *i.e.* we have set $\mathbf{Z} = \mathbf{I}$.

The software includes a five-bit control number set by the user which operates as follows:

bit	set	clear
1	σ_r read from data file	σ_r set to the same value for all reflections
2	σ_t read from data file	σ_t set to the same value for all reflections
4	σ_t multiplied by $\|F_o\|$ from data file	σ_t unchanged
8	σ_t multiplied by $(\sqrt{1-m^2}/m)$ with m taken from data file (m = figure of merit)	σ_t unchanged
16	φ read from data file	$\varphi = \alpha$

so that, for example, if this control number is zero and $\sigma_t^{-1} = 0$ then the phase of F_o is treated as being equal to the phase of F_c and tangential errors are given no weight. This is conventional $|F|$ refinement in which the constant σ_r corresponds to $\sigma(I)/2\sqrt{I}$, which approximately models errors arising from counting statistics. Setting the control number to 1 refines against $|F|$ but with individual weights for the reflections. Method 29 sets φ and σ_r from the data file and σ_t proportional to $|F_o|$ times the tangent of the uncertainty in φ.

4. Method

The radii a_{jk} in (16) are the radii at which the density of the kth component of a stationary atom has fallen to $e^{-\pi}$, about 1/20th, of its central value. These have been set to the values given in Table 1, which, together with the stated electron counts, reproduce the figures given for atomic scattering factors in Table 3.3.1A of *International Tables for X-ray Crystallography* (1968) with a r.m.s. error in the range $0 \leq |\mathbf{s}| \leq 1 \cdot 0$ Å^{-1} given in the bottom row.

The internal modes used were like the ten lowest-frequency modes of Levitt, Sander & Stern (1985) and were kindly supplied by M. Levitt. The calculation of the modes used here, however, included all hydrogen atoms, whereas those of the 1985 paper included only a subset of the polar hydrogens. These authors give in their paper reason to believe that most of the molecular motion is attributable to the eight lowest-frequency modes.

The external modes used were of two kinds, translational and rotational. The translational modes consisted of uniform displacements of all atoms in one molecule along each of the six ⟨110⟩ directions (in a cubic Cartesian coordinate system, rather than the crystallographic ⟨110⟩). These six were chosen because they are a minimum set which together may synthesize an overall anisotropic temperature factor of arbitrary orientation. The rotational modes provided for libration about each of

the three principal axes of inertia through the centroid of each molecule. No provision was made for screw motions.

The data set used to define the BPTI molecule on input was constructed by merging data from two sources. One of these is the data set 5PTI from the Protein Data Bank (Bernstein *et al.*, 1977) deposited by Huber and Wlodawer following their combined X-ray and neutron refinement of form II of the bovine pancreatic trypsin inhibitor (Wlodawer, Walter, Huber & Sjölin, 1984). These authors report that a restrained refinement of anisotropic thermal parameters was performed by the method of Hendrickson & Konnert (1981) in which thermal parameters of nearest and next-nearest neighbours are restrained to be similar. Nevertheless 5PTI contains only isotropic B factors, presumably extracted from the anisotropic set by the method of Hamilton (1959). The existence of these restraints means that the 892 B factors in group 1 are not entirely independent, but their averages, a residue at a time, must be little affected by this.

The 5PTI data set contains coordinates, occupancies and isotropic temperature factors for 1105 atoms consisting of

group 1	892 atoms, including H and D, for the protein
group 2	17 atoms representing alternative conformations for the side chains of residues 7 and 52
group 3	194 atoms representing 63 molecules of D_2O and one phosphate ion
group 4	2 atoms given zero weight.

The second data set, supplied by M. Levitt, contains coordinates and modal displacement matrices for the ten lowest-frequency modes for all atoms in group 1 and no others. Both data sets use Cartesian coordinates in Å rather than crystallographic fractional coordinates and all the internal working of the current program [except for the FFT (fast Fourier transform) operations] uses these too.

For each atom in the Levitt data set, an atom, matching by position, was found in 5PTI and the corresponding entry in the merged data set then consisted of the coordinates from the Levitt source, occupancies from 5PTI, negated temperature factors from 5PTI, the ten modal displacement vectors from Levitt and the nine displacement vectors for the external modes as already described. The Levitt coordinates have been used because these are the ones used by him when calculating the modal displacements, so that the model used is self consistent. They differ from the 5PTI coordinates by 0·05315 Å r.m.s. distance over the atoms in group 1. This difference of coordinates, while it may slightly raise the R factor, has only a second-order effect on an optimization in which only scale and thermal parameters are refined since derivatives of F_c or of $\rho(\mathbf{r})$ with respect to position are orthogonal to those with respect to scale or thermal parameters. The 5PTI temperature factors were retained for later comparison purposes but were negated to indicate that such atoms are to have their thermal parameters handled by the methods of this paper.

Atoms in groups 2 and 3 were then transferred from 5PTI to the merged data set using coordinates, occupancies and temperature factors all from this source. They were also given vanishing modal displacement vectors. These atoms then contribute to F_c in the manner attributed to them by Wlodawer *et al.* but contribute nothing to $\partial F_c/\partial p$. They do, however, participate in the determination of the overall scale factor, S. Atoms in group 2 have partners in group 1 with which their occupancies sum to unity. Occupancies are treated here as non-adjustable multipliers on the Z values.

Of the two atoms in group 4 one is listed as 'unknown, probably potassium' in 5PTI but is given an atom type identifier of zero. This atom has been omitted from the working owing to uncertainty concerning the appropriate stationary-atom scattering factor to use. The other is a software artefact at the origin (mine), and is zero weighted.

The reciprocal-space data set used is R5PTISFX from the Protein Data Bank, also deposited by Wlodawer and Huber. This data set contains 17555 independent reflections in the resolution range 6–1 Å, but does not contain individual values of expected standard deviations for the structure factors. Consequently method 0 has been used with σ_t^{-1} set to zero (see above).

Operationally, the software consists of several free-standing Fortran programs (available on request) linked together into a cycle at the command procedure level (VAX/VMS DCL) and run on a VAX 8600. This has permitted standard packages (*e.g.* for FFT) and standard file structures to be used consistent with the practices of the Collaborative Computing Project 4 (CCP4) of the UK Science and Engineering Research Council. Specifically, in scanning the molecule to construct maps of $\rho(\mathbf{r})$ and $\partial\rho(\mathbf{r})/\partial p_p$, function values for all 20 maps at a single grid point are stored in consecutive locations, for reasons of efficiency at this stage. The resulting file has 20 interleaved maps which then need to be separated before transformation. On transforming each map a 'labelled column format' (LCF) file is constructed, a column each for $|G|$ and β being added for each map, so that the matrix of derivatives for the least-squares stage then consists of a LCF file of some 48 columns.

On output, the matrix $\mathbf{NPN}^T$, in Å^2, is given for each atom together with the eigenvectors of each such matrix scaled by the square root of the corresponding eigenvalue.

These vectors then represent, in Å, the r.m.s. displacements in the three principal directions for each atom. Furthermore, the square root of the arithmetic mean eigenvalue then gives the radius of an equivalent spherical distribution having the same r.m.s. displacement (Hamil-

ton, 1959), and this may then be compared with the isotropic r.m.s. displacement (in any one direction) given by $(B/8\pi^2)^{1/2}$ using the B values in 5PTI. Average and r.m.s. values of both these quantities were obtained on a residue-by-residue basis for comparative plots, and correlation coefficients between them were calculated on an atom-by-atom basis.

For any group of positive quantities with mean $\bar{x}$ and variance σ^2 the ratio of their geometric mean to their arithmetic mean is $\exp(-\sigma^2/2\bar{x}^2)$ to a good approximation for small values of the third and higher moments of their distribution. Thus, if we obtain a second equivalent radius as the square root of the geometric mean eigenvalue, which corresponds to the radius of the sphere of equal volume, a plot of both these equivalent radii as a function of residue number gives an indication of the scatter of the eigenvalues, and thus of the anisotropy.

5. Results

Fig. 1 summarizes the results obtained for the protein atoms (group 1) and compares these with the isotropic temperature factors given by Wlodawer *et al.* (1984), in 5PTI. The broken curve is the quantity $(B/8\pi^2)^{1/2}$ from 5PTI, averaged over the atoms of each residue and plotted as a function of residue number. The two highest continuous curves are the two forms of equivalent radii described above, similarly averaged over the atoms of each residue. It is evident from these that the 19 thermal parameters used in this work provide a description of the thermal motion which closely matches the description given by 892 isotropic parameters. This pair of curves also shows, as is to be expected, that the greatest anisotropy occurs where the apparent motion is greatest, especially at residue 39 and the C terminus, at which points spheres of equal volume compare with the isotropic B values more closely than do the spheres of equal r.m.s. displacement. The correlation coefficients between the radii derived from the isotropic B values and the equivalent radii found here are 0·720 for the spheres of equal r.m.s. displacement and 0·783 for the spheres of equal volume, these figures being derived from the individual atomic values, not from the residue averages.

Having determined all 19 diagonal elements of **P** one may then enquire how much of the apparent motion arises from each source by replacing $\mathbf{NPN}^T$ for each atom by such a matrix in which only selected elements of **P** are non-zero, these being given the values which arise when all 19 are determined together. When this technique is used to isolate the contribution from the six overall trans-

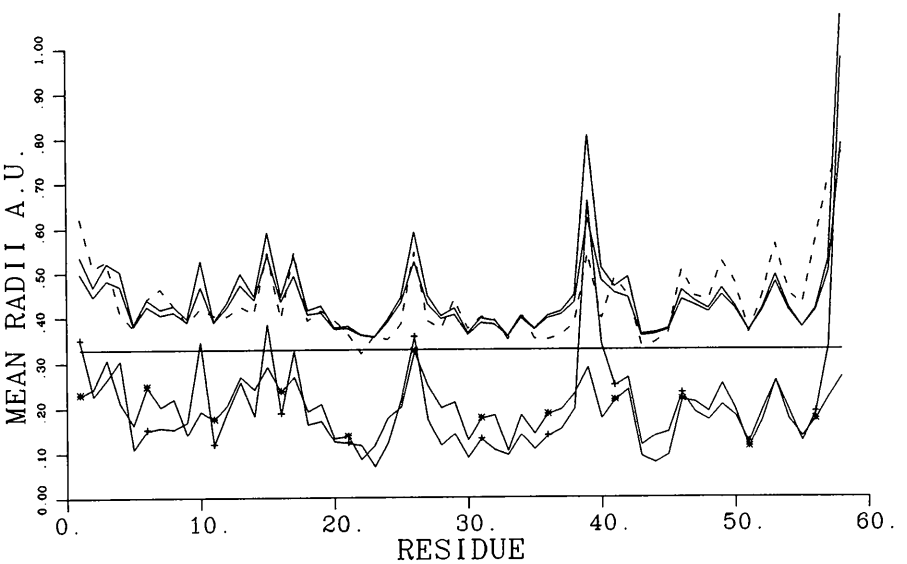

Fig. 1. Average apparent atomic radius plotted against residue number. The broken curve consists of the quantity $(B/8\pi^2)^{1/2}$ derived from Wlodawer and Huber's B values given in 5PTI, averaged over the atoms in each residue (group 1), and plotted against residue number. The top two solid curves are radii of isotropic equivalents to the anisotropic thermal parameters produced in this work. The upper one of the two is the radius of a spherical distribution of equal r.m.s. displacement and the lower one is the radius of a sphere of equal volume. These two curves diverge where the distribution $\mathbf{NPN}^T$ is strongly anisotropic and coincide where it is spherical. These curves are likewise averaged over the atoms of each residue and should be comparable to the broken curve. The remaining solid lines are radii of spheres of equal r.m.s. displacement. Of these the straight line is the contribution attributed to the six translational modes in the full optimization (19 parameters plus scale factor) and thus represents an overall anisotropic B factor attributable to lattice motion or lattice disorder, and is a property of the crystal rather than an intrinsic property of the molecule. It is evidently the largest single contribution. The curve marked with asterisks is the contribution attributed to the three libration modes and thus represents the orientational motion or disorder which, likewise, is not an intrinsic property of the molecule, being controlled by the precision with which each molecule is oriented on its site. The remaining curve, marked with + signs, is the contribution attributed to the ten lowest frequency modes of Levitt, Sander & Stern, and is thus a molecular property representing contributions arising from motions internal to each molecule.

lational modes (lattice modes) the horizontal line at about 0·33 Å results. Taken literally, this means that if one extrapolates the crystal lattice from the vicinity of one unit cell to some part of the crystal which is remote but still within the X-ray coherence volume, then unit cells in the remote region are expected to be randomly displaced relative to the extrapolation by some 0·33 Å on each coordinate, though, as we show below, this figure represents an upper limit. This figure itself is made up of a combination of long-period motions and time- and space-dependent fluctuations in cell size arising from the presence of a population of water molecules in thermal motion between the protein molecules. Variations in the details of intermolecular contacts also contribute to this figure.

The two remaining curves in the diagram are both graphs of the radii of spheres of equal r.m.s. displacement, the one marked with + signs being the contribution from the ten modes of M. Levitt, and the remaining curve (marked with asterisks) the contribution from the three libration modes. The remarkable similarity of these two curves arises because rotational modes (by definition) and internal modes (because they have zero net momentum) tend to have their largest motions on the surface of the molecule; thus the variation in the two curves tends to reflect the distinction between internal and surface residues.

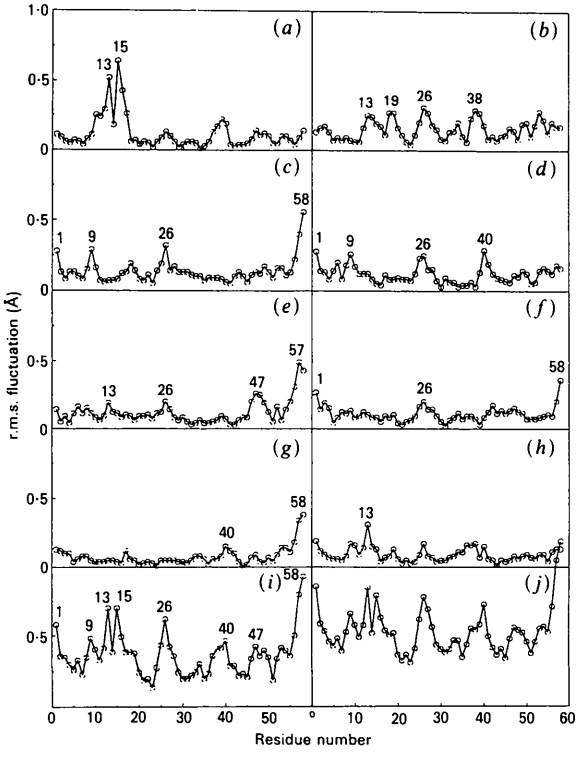

Fig. 2. Root mean square displacements given by Levitt, Sander & Stern. Each graph is the r.m.s. displacement of α-carbon atoms plotted as a function of residue number for the eight lowest-frequency modes in (*a*) to (*h*). These are calculated to correspond to a BPTI molecule *in vacuo* at 300 K. (*i*) shows the r.m.s. displacement resulting from the combined displacement of the first eight modes, and (*j*) likewise from all 208 modes. All the modes are derived using only the torsion angles in single bonds as variables, bond lengths and angles being treated as constants. Figure reproduced with permission from Fig. 7 of Levitt, Sander & Stern (1985).

It is evident from Fig. 1 that lattice effects which yield an overall anisotropic temperature factor provide the largest of the three contributions to the observed *B* factors. It is also apparent that the level of displacement calculated by Levitt (Figs. 2*i* and *j*) (which should be higher by a factor $\sqrt{3}$ than the corresponding curve in Fig. 1), becomes comparable with the component attributed to the internal modes in Fig. 1, but could not otherwise be reconciled with the complete *B* values.

It is important that this separation of the apparent displacements into inter- and intramolecular effects be estimated accurately since the intramolecular contribution can be expected to be similar in solution and to have some bearing on the reaction kinetics of BPTI. To some extent the apparent distribution as between these two types of contribution must be a function of the number and types of modes offered to the optimization. Fig. 3 shows the effect of varying this number. Independent optimizations were done for each of 6, 9, 14 and 19 modes, in each of which the scale factor was also optimized. These groupings were respectively the six translations, these plus the three librations, these plus the five lowest-frequency internal modes, and these plus the next five internal modes.

Both types of mean equivalent radius, averaged over all the atoms of group 1, are plotted using circles and squares. The upper pair of curves uses all the modes in the optimization concerned and rises gradually to bracket the value 0·443 Å, which is the mean of $(B/8\pi^2)^{1/2}$ derived from the values of *B* given by Wlodawer *et al.* for the group 1 atoms in the data set 5PTI. This is the value to be expected from the modal method if the number of modes employed were sufficient to reproduce all the detailed fluctuations of *B* factor which are present in 5PTI. The lower curve is the contribution of the six translational modes alone in each of these optimizations and is remarkably stable when nine or more modes are involved, suggesting that the lattice contribution to the apparent motion is already quite well estimated. This fact must be interpreted with some caution however, because internal motion such as may be characterized by modes 11 to 208 of Levitt, Sander & Stern (1985), which are not modelled by the present optimization, if present in the molecule must be represented by whatever parameters are available to the optimization. Careful comparison of Figs. 2(*i*) and (*j*) shows that the change in r.m.s. displacement of α-carbon atoms occasioned by adding modes 9 to 208 to modes 1 to 8 is calculated to be to add 0·17 Å overall with little other change, and this figure may be used to estimate a lower limit for the lattice contribution in the present work. In Levitt's notation we may write

$$[\sigma_{\alpha i}^{1\text{ to }8}]^2 + [\sigma_{\alpha i}^{9\text{ to }208}]^2 = [\sigma_{\alpha i}^{1\text{ to }208}]^2 \approx [\sigma_{\alpha i}^{1\text{ to }8} + \Delta\sigma]^2 \quad (34)$$

$$\sigma_{\alpha i}^{9\text{ to }208} \approx \sqrt{2\sigma_{\alpha i}^{1\text{ to }8}\Delta\sigma + (\Delta\sigma)^2} = 0\cdot 39\text{Å} \quad (35)$$

using the mean of $\sigma_{\alpha i}^{1\text{ to }8}$ from Fig. 2(*i*), which is 0·36 Å, in place of $\sigma_{\alpha i}^{1\text{ to }8}$ in (35). Supposing that internal motion of this order arises from modes 11 to 208 and that this is subsumed within what has been modelled here as lattice displacements, then

$$[\sigma_{\alpha i}^{11\text{ to }208}]^2 + [\sigma_{\alpha i}^{\text{lattice}}]^2 = 3 \times 0 \cdot 33^2 \text{Å}^2$$
$$\sigma_{\alpha i}^{\text{lattice}} = 0 \cdot 42 \text{Å}$$
$$\text{or } 0 \cdot 24 \text{ Å on each coordinate} \tag{36}$$

Thus the positional uncertainty arising from perturbations of the lattice is estimated to lie between 0·24 and 0·33 Å on each coordinate. This result may be compared with the result of Edwards *et al.* (1990) for ribonuclease-A based on measurement of the velocity of sound of $\langle u^2 \rangle = 0 \cdot 06 \pm 0 \cdot 03 \text{Å}^2$ which is a positional uncertainty of $\sqrt{0 \cdot 06} = 0 \cdot 24$ Å on each coordinate. The present results are therefore compatible with their results and can accommodate also a contribution from static disorder which, if present, is excluded from their estimate. The observed upper limit figure of 0·33 Å r.m.s. (or 0·11 Å^2 mean square) is closely similar to the value found for lysozyme by Sternberg, Grace & Phillips (1979) who report values of 0·115 and 0·110 Å^2 for the corresponding quantity in the **TLS** and **TL** models respectively.

The r.m.s. values of the three librational angles are 1·99, 0·92 and 1·07° of which the largest is associated with the long axis of the molecule running approximately from Lys 15 to Ala 58. These figures are similar to those found for ribonuclease-A by Howlin, Moss & Harris (1989), 1·4° r.m.s., and for lysozyme by Sternberg, Grace & Phillips (1979) who report r.m.s. values of 2·6, 1·6 and 0·5 °.

The diagonal elements of **P** for the internal modes (when the **N** matrices are scaled to correspond to 300 K) are 0·766, 0·000, 1·302, 0·212, 0·070, 0·772, 0·000, 0·341, 0·828 and 0·837, which have mean 0·513 and standard deviation 0·44. Ideally, the Boltzmannian values for these figures would all be 0·5. Estimates of expected standard deviations for the individual elements are not offered because the data set R5PTISFX does not include e.s.d.s on the $|F_o|$ values, and because they are highly correlated observationally (although uncorrelated temporally), as already remarked after (33).

The overall *R* factor and scale factor are also plotted in Fig. 3, together with their limiting values which are obtained by optimizing *only* the scale factor to be applied to 5PTI, and measuring the *R* factor with the resulting value and the given *B* values. This one-parameter optimization was included to obtain strict comparability with the results of the modal method, which comparisons might otherwise have been influenced by minor differences between atomic scattering factors, cut-off limits in modelling the density, or other minor differences of technique between the current work and that of Wlodawer, Walter, Huber & Sjölin (1984).

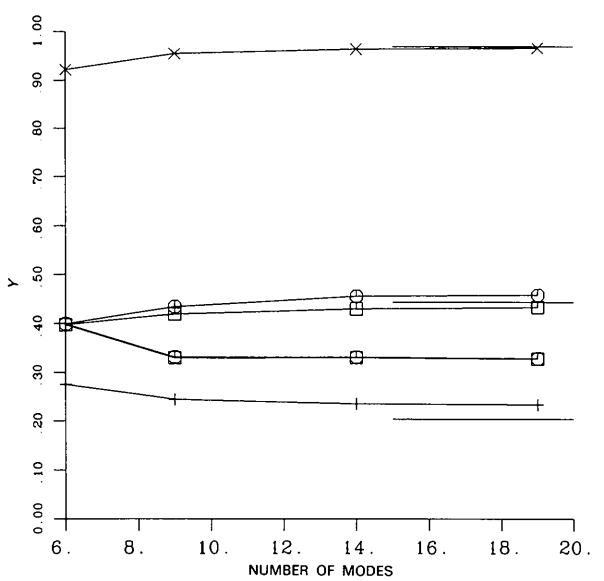

Fig. 3. The dependence of apparent radii, scale factor and *R* factor on the number of modes employed. Modes 1 to 6 are the translational modes, 7 to 9 the librational ones and 10 to 19 the internal modes of Levitt, Sander & Stern. The curves marked with circles and squares are respectively equivalent radii of equal r.m.s. displacement and of equal volume, averaged over all 892 atoms in group 1. The two upper curves so marked are based on all the modes used in each optimization, and the lower curve (the two types of curve coincide) is the contribution of the six translational modes, which is little affected by the introduction of the internal modes to the system. For these curves the vertical scale is in Å and the horizontal bar is the equivalent mean radius of Wlodawer and Huber. For a discussion of these curves see text. The top curve is the scale factor, the horizontal bar being the value obtained using the isotropic *B* factors of Huber and Wlodawer. The bottom curve is the *R* factor, as a fraction, the horizontal bar being the value obtained using Huber and Wlodawer's 892 *B* factors and an independently measured scale factor.

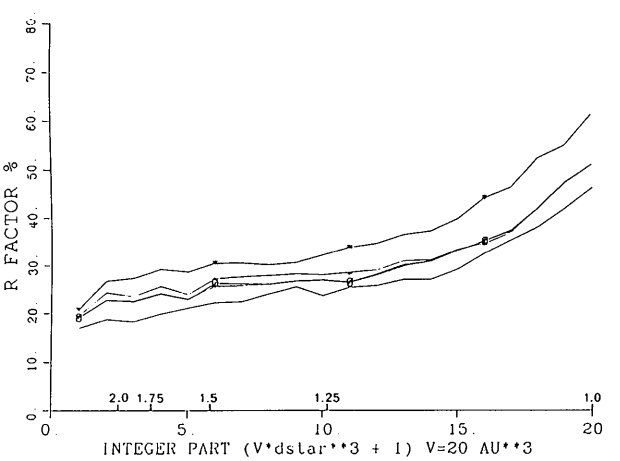

Fig. 4. Conventional *R* factors plotted against $V[(2 \sin \theta)/\lambda]^3$ with V=20Å^3. The upper abscissa scale is $[(2 \sin \theta)/\lambda]^{-1}$ in Å. In descending order these curves are for 6, 9, 14, 19 and 892 thermal parameters.

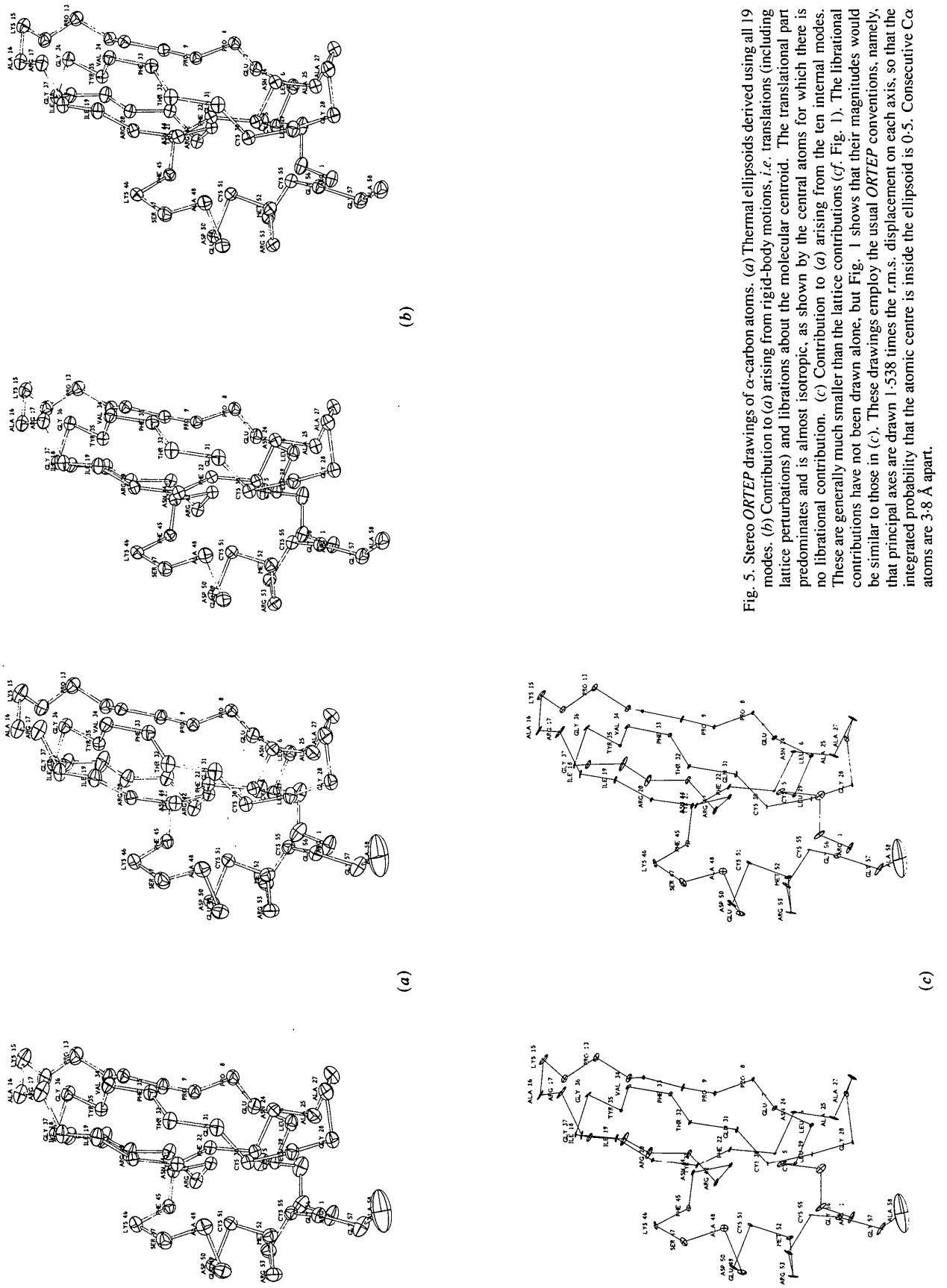

Fig. 5. Stereo *ORTEP* drawings of α-carbon atoms. (*a*) Thermal ellipsoids derived using all 19 modes. (*b*) Contribution to (*a*) arising from rigid-body motions, *i.e.* translations (including lattice perturbations) and librations about the molecular centroid. The translational part predominates and is almost isotropic, as shown by the central atoms for which there is no librational contribution. (*c*) Contribution to (*a*) arising from the ten internal modes. These are generally much smaller than the lattice contributions (*cf.* Fig. 1). The librational contributions have not been drawn alone, but Fig. 1 shows that their magnitudes would be similar to those in (*c*). These drawings employ the usual *ORTEP* conventions, namely, that principal axes are drawn 1·538 times the r.m.s. displacement on each axis, so that the integrated probability that the atomic centre is inside the ellipsoid is 0·5. Consecutive Cα atoms are 3·8 Å apart.

Fig. 4 shows the R factor plotted against d^{*3}. Here the reflections were sorted into 20 spherical shells of equal volume, the outer limit of the outermost shell corresponding to $(2 \sin\theta)/\lambda = 1 \cdot 0$ Å^{-1}. Both these figures indicate that the three librational modes used are highly effective in reducing the R factor, suggesting that the variation in B between the surface and the interior of the molecule is the predominant systematic variation of this quantity. No optimization has been done in which the model included the internal modes and the translations, to the exclusion of the librations.

Finally, in Fig. 5 we show stereo pairs of *ORTEP* drawings (Johnson, 1976) of the α-carbon atoms alone. It is noticeable that internal and bridged residues, such as Phe 22 and Cys 55, have closely similar almost spherical distributions which are dominated by the lattice contribution. Other regions, especially the termini, show marked elongations with a tendency for the major axis of each ellipsoid to be transverse to its position vector relative to the molecular centre. The portion of chain from Gly 37 to Arg 42 also shows substantial anisotropy with the elongation largely transverse to the length of the chain. Lateral movements of this portion correspond to soft deformations of the molecule and to displacements as between form I and form II of this molecule (Wlodawer, Deisenhofer & Huber, 1987).

References

BERNSTEIN, F.C., KOETZLE, T.F., WILLIAMS, G.J.B., MEYER, E.F. JR, BRICE, M.D., RODGERS, J.R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

DIAMOND, R. (1976). In *Crystallographic Computing Techniques*, edited by F.R. AHMED, pp. 291–301. Munksgaard: Copenhagen.

EDWARDS, C., PALMER, S.B., EMSLEY, P., HELLIWELL, J.R., GLOVER, I.D., HARRIS, G.W. & MOSS, D.S. (1990). *Acta Cryst.* A**46**, 315–320.

HAMILTON, W.C. (1959). *Acta Cryst.* **12**, 609–610.

HENDRICKSON, W.A. & KONNERT, J.H. (1981). In *Biomolecular Structure, Conformation, Function and Evolution*, edited by R. SRINIVASAN, pp. 43–58. Oxford: Pergamon.

HOWLIN, B., MOSS, D.S. & HARRIS, G.W. (1989). *Acta Cryst.* A**45**, 851–861.

International Tables for X-ray Crystallography (1968). Vol. III. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)

JOHNSON, C.K. (1976). *ORTEP*II. *A Fortran Thermal Ellipsoid Plot Program for Crystal Structure Illustrations*. Report ORNL-5138, Oak Ridge National Laboratory, Tennessee, USA.

KONNERT, J.H. & HENDRICKSON, W.A (1980). *Acta Cryst.* A**36**, 344–350.

LEVITT, M., SANDER, C. & STERN, P.S. (1985). *J. Mol. Biol.* **181**, 423–447.

SCHOMAKER, V. & TRUEBLOOD, K.N. (1968). *Acta Cryst.* B**24**, 63–76.

STERNBERG, M.J.E., GRACE, D.E.P. & PHILLIPS, D.C. (1979). *J. Mol. Biol.* **130**, 231–253.

WLODAWER, A., DEISENHOFER, J. & HUBER, R. (1987). *J. Mol. Biol.* **193**, 145–156.

WLODAWER, A., WALTER, J., HUBER, R. & SJÖLIN, L. (1984). *J. Mol. Biol.* **180**, 301–329.

Acta Cryst. (1990). A**46**, 435–437

Temperature Factor of Silicon by Powder Neutron Diffraction

BY ZHANG BAISHENG, YANG JILIAN, JIN LAN AND YE CHUNTANG

Institute of Atomic Energy, PO Box 275 (30), *Beijing, People's Republic of China*

AND J. BASHIR, N. M. BUTT,* M. SIDDIQUE, M. ARSHED AND Q. H. KHAN

Pakistan Institute of Nuclear Science and Technology, PO Nilore, Islamabad, Pakistan

(*Received* 14 *September* 1989; *accepted* 30 *November* 1989)

Abstract

The temperature factor of silicon has been determined by the powder neutron diffraction technique employing a double-axis neutron diffractometer. A neutron wavelength of 1·184 Å was used in the experiment. The sample used was a fine powder of silicon of purity 99·999%. The correction to the observed intensities due to thermal diffuse scattering (TDS) was not applied as the neutron velocity of 3·34 km s^{-1} (corresponding to neutron wavelength of 1·184 Å) is less than the minimum velocity of sound in this crystal. The B value obtained from these experiments was found to be 0·45 (2) Å^2, corresponding to a mean-square vibrational amplitude of 0·017 (2) Å^2 and to a Debye temperature of 531 (11) K at the sample temperature of 284 K at which the experiment was performed.

Introduction

Silicon is an important element of commercial interest and accurate measurement of its physical properties is desirable. One such physical parameter is the total

* To whom all correspondence should be addressed.

0108-7673/90/060435-03$03.00