Journal of
Applied
Crystallography

ISSN 0021-8898

On the handling of atomic anisotropic displacement parameters

R. W. Grosse-Kunstleve and P. D. Adams

Journal of
**Applied
Crystallography**
ISSN 0021-8898

On the handling of atomic anisotropic displacement parameters

R. W. Grosse-Kunstleve* and P. D. Adams

Lawrence Berkeley National Laboratory, One Cyclotron Road, Mail Stop 4-230, Berkeley, California 94720, USA. Correspondence e-mail: rwgrosse-kunstleve@lbl.gov

Received 10 December 2001
Accepted 9 May 2002

A number of conventions for the parameterization of atomic anisotropic displacements are used in the literature and in crystallographic programs. Here we summarize the commonly used conventions, with a special emphasis on their application in macromolecular crystallography. We then describe a new software toolbox for the handling of the various parameterizations of atomic anisotropic displacements and their interconversion. All algorithms are integrated into the freely available *Computational Crystallography Toolbox*.

1. Introduction

When high-resolution diffraction data are available, crystal structures are often modeled using atomic anisotropic displacement parameters (ADPs) to describe thermal motion and possible static displacive disorder. In the field of small-molecule crystallography, ADPs have been routinely used for decades (Trueblood *et al.*, 1996). Until recently, the use of ADPs in the refinement of macromolecular structures has been rare. However, as a result of many improvements in data collection technology, high-resolution data that enable the refinement of ADPs are increasingly becoming available for macromolecules. This is reflected by a rapid growth in the number of structures with ADPs deposited in the Protein Data Bank (PDB) (Berman *et al.*, 2000) (Fig. 1).

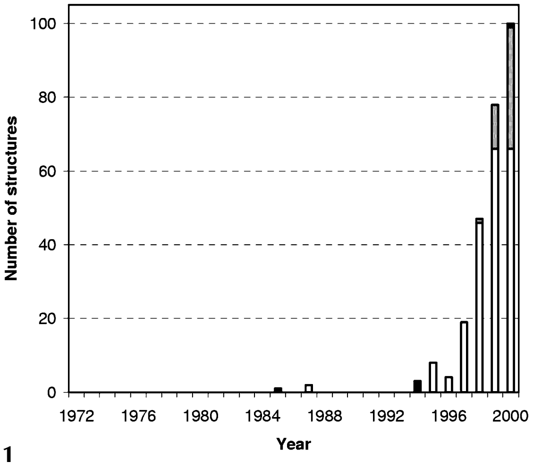

Figure 1
Histogram of structures that were deposited in the Protein Data Bank (PDB) with ANISOU cards, grouped by year, based on the PDB holdings on 16 October 2001. The dates used for the histogram are the days of deposition as found on the PDB HEADER card. The total number of structures with ANISOU cards is 306. 237 structures were refined with *SHELX* (white bars) (Sheldrick & Schneider, 1997), 64 structures with *REFMAC* (gray bars) (Murshudov *et al.*, 1997), and five structures with three other programs (black bars) (see the references archived in the PDB under the access codes 7bna, 1etl, 1etm, 1etn and 1ejg).

As we show in §2, a number of conventions for the parameterization of anisotropic displacements are currently in use. These different conventions present a significant potential for confusion. For example, the ADPs found in files that follow the PDB format (http://www.rcsb.org/pdb/info.html) follow a different convention with respect to the ADPs found in files that follow the mmCIF format (http://pdb.rutgers.edu/mmcif/). If this is not taken into account, there exists a possibility that the ADPs can be misinterpreted, leading to incorrect analysis of a structure. Therefore, a library for the conversion between the different parameterizations is a valuable tool. We have implemented such a library by adding the ADP toolbox (*adptbx*) to the *Computational Crystallography Toolbox* (*cctbx*) (Grosse-Kunstleve *et al.*, 2002). In addition to the conversions, the library facilitates the computation of Debye–Waller factors, the handling of symmetry restrictions, and the determination of the eigenvalues and eigenvectors of anisotropic displacement ellipsoids. In the following section, we will summarize the commonly used conventions for ADPs in the literature, computer programs and databases. This is followed by a description of the *adptbx*. Further documentation is available online (http://cctbx.sourceforge.net/).

2. Commonly used conventions for the parameterization of atomic anisotropic displacement parameters

The mean-square displacements that define the probability density functions of atomic displacements are commonly parameterized as a trivariate Gaussian. The effect of the atomic displacements enters into the structure-factor calculation as the Debye–Waller factor $T(\mathbf{h})$, where $\mathbf{h}$ is a column vector with the Miller indices of a Bragg reflection. The fundamental expression for $T(\mathbf{h})$ is

$$T(\mathbf{h}) = \exp\left[-2\pi^2\langle(\mathbf{h}\cdot\mathbf{u})^2\rangle\right], \qquad (1)$$

where $\mathbf{u}$ is a row vector with the components of the displacement vector (Trueblood *et al.*, 1996, equations 14–22 therein). Equation (1) takes on different forms depending on the basis vectors to which the diffraction and displacement vectors are referred. If the diffraction vectors are referred to the basis of the crystallographic reciprocal lattice, and the displacement vector to the crystallographic direct lattice, equation (1) takes on the form

$$T(\mathbf{h}) = \exp(-2\pi^2 \mathbf{h}^t \mathbf{U}^* \mathbf{h}), \tag{2}$$

where $\mathbf{h}^t$ is the transpose of $\mathbf{h}$ (*i.e.* a row vector), and $\mathbf{U}^*$ is a symmetric second-rank tensor, typically represented as a symmetric 3×3 matrix of real numbers. This definition of $\mathbf{U}^*$ was established by Giacovazzo (1992) and Murshudov *et al.* (1999). [Bricogne (2001) uses the notation $\mathbf{Q}$ for the same tensor.] The elements of $\mathbf{U}^*$ are dimensionless mean-square displacements, analogous to fractional coordinates. This definition of ADPs can be used directly in the structure-factor calculation (without explicitly involving the unit-cell parameters) and is the most suitable parameterization for establishing compatibility with the site-symmetry for special positions (see §3.3).

All other commonly used conventions for the parameterization of anisotropic displacements are most conveniently defined through tensor transformations. For example, let $\mathbf{A}$ be an orthogonalization matrix that transforms the fractional coordinates $\mathbf{x}_{\rm frac}$ with respect to a crystallographic basis system to coordinates $\mathbf{x}_{\rm cart}$ with respect to a Cartesian basis. Application of the transformation rules listed by Giacovazzo (1992, section 2.E therein) leads to

$$\mathbf{U}_{\rm cart} = \mathbf{A}\mathbf{U}^*\mathbf{A}^t, \tag{3a}$$

$$\mathbf{U}^* = \mathbf{A}^{-1}\mathbf{U}_{\rm cart}(\mathbf{A}^{-1})^t, \tag{3b}$$

where $\mathbf{A}^{\rm t}$ is the transpose of $\mathbf{A}$, $\mathbf{A}^{-1}$ is the inverse of the system $\mathbf{A}$ [the elements of $\mathbf{A}$ have dimension (length)] and $\mathbf{U}_{\rm cart}$ is the anisotropic displacement tensor with respect to the Cartesian basis. The elements of $\mathbf{U}_{\rm cart}$ have dimension (length)^2. [$\mathbf{U}_{\rm cart}$ corresponds to U^C_{jl} of Trueblood *et al.* (1996).]

The ANISOU cards of PDB files define the elements of $\mathbf{U}_{\rm cart}$. The parameterization of anisotropic displacements with respect to a Cartesian basis system is useful because the eigenvalues of $\mathbf{U}_{\rm cart}$ are directly equivalent to the mean-square atomic displacements in the principal directions of the corresponding ellipsoid (see §3.4) and are also most suitable for generating graphical representations of such ellipsoids [as performed *e.g.* by the *ORTEP* program (Burnett & Johnson, 1996)].

In small-molecule crystallography, the dimensionless basis ($a^*\mathbf{a}$, $b^*\mathbf{b}$, $c^*\mathbf{c}$) is widely used for the parameterization of anisotropic displacements (Trueblood *et al.*, 1996, equations 23–25 therein; a^*, b^* and c^* are the lengths of the basis vectors of the crystallographic reciprocal lattice; $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ are the basis vectors of the crystallographic direct lattice). To our knowledge, this *ad hoc* basis system is only used for the parameterization of anisotropic displacements. It leads to the relations

$$\mathbf{U}_{\rm cif} = \mathbf{N}^{-1}\mathbf{U}^*(\mathbf{N}^{-1})^t, \tag{4a}$$

$$\mathbf{U}^* = \mathbf{N}\mathbf{U}_{\rm cif}\mathbf{N}^t, \tag{4b}$$

with

$$\mathbf{N} = \begin{pmatrix} a^* & 0 & 0 \\ 0 & b^* & 0 \\ 0 & 0 & c^* \end{pmatrix}. \tag{5}$$

The components of the tensor $\mathbf{U}_{\rm cif}$ have dimension (length)^2. [$\mathbf{U}_{\rm cif}$ corresponds to U^{jl} of Trueblood *et al.* (1996)]. This convention for ADPs was adopted in the definitions of both the CIF (http://www.iucr.org/iucr-top/cif/) and the mmCIF dictionaries. For clarity, in this paper and in the *adptbx*, the parameters according to the CIF definition (dictionary item `aniso_U`) are consistently labeled as $\mathbf{U}_{\rm cif}$, and the parameters according to the PDB convention (ANISOU cards) are consistently labeled as $\mathbf{U}_{\rm cart}$. Note that substituting equation (4*b*) into equation (2) results in the lengthy expression found in the CIF dictionary and in many program descriptions [*e.g.* the *SHELX* manual (Sheldrick & Schneider, 1997)].

For unit cells with orthogonal basis vectors, $\mathbf{U}_{\rm cif}$ is exactly equivalent to $\mathbf{U}_{\rm cart}$. Unfortunately, for unit cells with angles other than 90°, the mean-square displacements are convoluted with the metric of the crystallographic basis. To determine the principal mean-square displacements for the general case, it is therefore necessary to convert $\mathbf{U}_{\rm cif}$ to $\mathbf{U}_{\rm cart}$. The conversion laws are easily obtained by combining equations (3) and (4).

The differences in the parameterizations introduced so far are exclusively in the choice of the basis systems. Alternative parameterizations found in the literature are convoluted with factors of π. The first example is the definition

$$\boldsymbol{\beta} = 2\pi^2\mathbf{U}^* \tag{6}$$

which is used by the *SIR* program (Burla *et al.*, 2000) [and corresponds to β^{jl} of Trueblood *et al.* (1996)]. The second example is the definition

$$B = 8\pi^2 U, \tag{7}$$

which is used in the definition of the CIF and mmCIF dictionaries [the factor $8\pi^2$ appears in the expression for the Debye–Waller factor given isotropic displacement parameters (see Trueblood *et al.*, 1996, equation 13 therein)]. The use of B is explicitly discouraged by Trueblood *et al.* (1996) and in the definitions of the CIF and mmCIF dictionaries. However, to be able to process input from sources that use B, support for this is included in the *adptbx*.

3. Overview of the *adptbx* and other related parts of the *cctbx*

3.1. Conversions

The *adptbx* supports the simple conversions between both isotropic and anisotropic 'U' and 'B' [equation (7)], independently of the basis system that is used. Further supported are all possible conversions between $\mathbf{U}^*$, $\mathbf{U}_{\rm cart}$, $\mathbf{U}_{\rm cif}$ and $\boldsymbol{\beta}$. This results in a total of 12 conversion functions. However, only the conversions according to equations (3) and (4) are generic implementations. The other conversions are implemented as nested function calls of the generic conversions. This minimizes the amount of source code, and the relations between the different parameterizations of the anisotropic displacements are also easy to follow. It should also be noted that modern optimizing compilers generate code that is as efficient at runtime as manually coded specific conversion functions.

Eight additional functions are provided for the conversions between the anisotropic $\mathbf{U}_{\rm cart}$, $\mathbf{U}^*$, $\mathbf{U}_{\rm cif}$ and $\boldsymbol{\beta}$, and the equivalent isotropic $U_{\rm iso}$. The isotropic equivalent of the anisotropic displacement tensor is defined as the mean of the mean-square displacements in the principal directions of the anisotropic displacement ellipsoid, which is equivalent to the mean of the eigenvalues. In the Cartesian basis system, the trace of the $\mathbf{U}_{\rm cart}$ tensor is equivalent to the sum of the eigenvalues (see *e.g.* Giacovazzo, 1992, p. 188 therein). Therefore, the equivalent $U_{\rm iso}$ is defined by the simple relation

$$U_{\rm iso} = \tfrac{1}{3}\,{\rm trace}\,(\mathbf{U}_{\rm cart}). \tag{8}$$

The conversion from $U_{\rm iso}$ to $\mathbf{U}_{\rm cart}$ is equally straightforward:

$$\mathbf{U}_{\rm cart} = \begin{pmatrix} U_{\rm iso} & 0 & 0 \\ 0 & U_{\rm iso} & 0 \\ 0 & 0 & U_{\rm iso} \end{pmatrix}. \tag{9}$$

The conversions between $\mathbf{U}^*$, $\mathbf{U}_{\rm uvrs}$ and $\boldsymbol{\beta}$, and the equivalent $U_{\rm iso}$ are implemented as the combination of conversions between *e.g.* $\mathbf{U}^*$ and $\mathbf{U}_{\rm cart}$, and $\mathbf{U}_{\rm cart}$ and $U_{\rm iso}$. In this way the complicated expressions for equivalent $U_{\rm iso}$ that are often found in the literature are avoided.

3.2. Computation of the Debye–Waller factor

Functions are provided for the computation of the Debye–Waller factor according to equation (1) given $\mathbf{U}^*$, $\boldsymbol{\beta}$, $\mathbf{U}_{\rm cart}$, $\mathbf{U}_{\rm cif}$, $U_{\rm iso}$ and $B_{\rm iso}$. Only two functions are generic implementations ($\mathbf{U}^*$, $B_{\rm iso}$). The others are implemented as a conversion followed by a call to the suitable generic function.

3.3. Handling of symmetry restrictions for special positions

For atoms on special positions, the *SiteSymmetry* class (Grosse-Kunstleve & Adams, 2002) in the *Space Group Toolbox* of the *cctbx* was extended to test whether a given anisotropic tensor $\mathbf{U}^*$ is compatible with the site symmetry. The anisotropic displacement ellipsoid must remain invariant under the application of each of the symmetry operations of the site-symmetry group. This leads to the condition

$$\mathbf{U}^* = \mathbf{R}\mathbf{U}^*\mathbf{R}^t, \tag{10}$$

where $\mathbf{R}$ is the rotation part of a given symmetry operation (see *e.g.* Giacovazzo, 1992).

A second function that was added to the *SiteSymmetry* class applies Wigner's theorem (Giacovazzo, 1992, p. 189 therein) to derive a tensor

$$\mathbf{U}^*_{\rm inv} = \tfrac{1}{n}\sum_{s=1}^{n}\mathbf{R}_s\mathbf{U}^*\mathbf{R}_s^t, \tag{11}$$

that is invariant under the n operations of the site-symmetry group even if $\mathbf{U}^*$ is not. This is useful to compensate for rounding errors arising from limited input precision, and possibly to compensate for approximations and rounding errors in refinement algorithms.

3.4. Determination of eigenvalues

The eigenvalues of $\mathbf{U}_{\rm cart}$ are directly equivalent to the mean-square atomic displacements in the principal directions of the anisotropic displacement ellipsoid. Since the anisotropic displacement tensor is a symmetric matrix, all eigenvalues are real. The eigenvalues λ are determined as the three real roots of the cubic equation

$$\left|\mathbf{U}_{\rm cart} - \lambda\mathbf{I}\right| = 0, \tag{12}$$

where $\mathbf{I}$ is the identity matrix. The solutions are obtained analytically using Cardan's formula (see *e.g.* Spiegel & Liu, 1998).

If any of the three eigenvalues is less than or equal to zero, the $\mathbf{U}_{\rm cart}$ tensor is not positive definite and $\mathbf{U}_{\rm cart}$ cannot be represented as an ellipsoid. Tests for this condition are included in the *adptbx*.

3.5. Determination of eigenvectors

The eigenvectors of $\mathbf{U}_{\rm cart}$ define the principal directions of the anisotropic displacement ellipsoid. Since the anisotropic displacement tensor is a symmetric matrix, the eigenvectors can be chosen to be orthonormal. The eigenvectors are determined using a simple method of successive approximations that is specific for 3 × 3 matrices, as outlined by Nye (1992, pp. 165–168 therein).

3.6. Software technology used

The *adptbx* is implemented in ISO C++ (International Standardization Organization *et al.*, 1998). This programming language supports object-oriented design, name-spaces, exception handling and parameterization of types (templates). These features facilitate the design of libraries that are reusable and maintainable. For example, all functions in the *adptbx* are parameterized (templated) by the floating-point type. It is therefore possible to choose between single precision and double precision. It is also possible to use multiple floating-point precisions simultaneously in the same program. Importantly, this is achieved without duplicating source code, and without impacting the runtime performance.

The *adptbx* may be used as a pure C++ library. However, for maximum flexibility, bindings to the high-level Python scripting language are provided. The benefits of this tight combination of a statically typed compiled language (C++) with a dynamically typed interpreted language (Python) are discussed in depth by Grosse-Kunstleve *et al.* (2002). A simple example script (`analyze_adp.py`) that demonstrates the use of the *adptbx* is provided online in the ‘examples’ directory of the *cctbx* (http://cctbx.sourceforge.net/).

4. Conclusion

We hope that the summary of commonly used conventions for the parameterization of atomic anisotropic displacements in §2 is a useful reference for crystallographers, in particular macromolecular crystallographers, who are unavoidably presented with at least two different parameterizations: the PDB convention ($\mathbf{U}_{\rm cart}$) and the mmCIF convention ($\mathbf{U}_{\rm cif}$).

We have introduced practical tools for the handling of ADPs, which are accessible through a scripting language (Python). This language was designed specifically for teaching programming concepts and the tools are therefore very approachable. Their integration into the *cctbx* creates a powerful resource for handling the diverse data representations used by crystallographic software. At the same time, their modular, highly reusable design, in combination with an Open Source license, makes them ideal for inclusion into large integrated software systems.

We would like to thank S. R. Hall, G. N. Murshudov and G. M. Sheldrick for responding to our questions. The American Mineralogist Crystal Structure Database (http://www.geo.arisona.edu/xtal-cgi/test) that is maintained by R. T. Downs was very helpful for the testing of the *adptbx*. Our work was funded in part by the US Department of Energy under Contract No. DE-AC03-76SF00098, and by NIH/NIGMS under grant number 1P01GM063210.

References

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). *Nucleic Acids Res.* **28**, 235–242.

Bricogne, G. (2001). *International Tables for Crystallography*, Vol. B, ch. 1.3. Dordrecht: Kluwer.

Burla, M. C., Camalli, M., Carrozzini, B., Cascarano, G. L., Giacovazzo, C., Polidori, G. & Spagna, R. (2000). *Acta Cryst.* A**56**, 451–457.

Burnett, M. N. & Johnson, C. K. (1996). *ORTEP-III*. Oak Ridge National Laboratory Report ORNL-6895 (http://www.ornl.gov/ortep/).

Giacovazzo, C. (1992). Editor. *Fundamentals of Crystallography*. IUCr/Oxford University Press.

Grosse-Kunstleve, R. W. & Adams, P. D. (2002). *Acta Cryst.* A**58**, 60–65.

Grosse-Kunstleve, R. W., Sauter, N. K., Moriarty, N. W. & Adams, P. D. (2002). *J. Appl. Cryst.* **35**, 126–136.

International Standardization Organization (ISO), International Electrotechnical Commission (IEC), American National Standards Institute (ANSI) & Information Technology Industry Council (ITI) (1998). *International Standard ISO/IEC 14882*, 1st ed., Information Technology Industry Council, 1250 Eye Street NW, Washington, DC 20005 (also available at http://webstore.ansi.org/).

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S. & Dodson, E. J. (1999). *Acta Cryst.* D**55**, 247–255.

Nye, J. F. (1992). *Physical Properties of Crystals.* Oxford: Claredon Press.

Sheldrick, G. M. & Schneider, T. R. (1997). *Methods Enzymol.* **277**, 319–343.

Spiegel, M. R. & Liu, J. (1998). *Mathematical Handbook of Formulas and Tables*. New York: McGraw-Hill.

Trueblood, K. N., Bürgi, H.-B., Burzlaff, H., Dunitz, J. D., Gramaccioli, C. M., Schulz, H. H., Shmueli, U. & Abrahams, S. C. (1996). *Acta Cryst.* A**52**, 770–781. (Also available at http://www.iucr.org/iucr-top/commun/cnom/adp/finrep/finrep.html.)