



An enhanced set of displacement parameter restraints in *CRYSTALS*

Pascal Parois, James Arnold and Richard Cooper

J. Appl. Cryst. (2018). **51**, 1059–1068



IUCr Journals
CRYSTALLOGRAPHY JOURNALS ONLINE

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site or institutional repository provided that this cover page is retained. Republication of this article or its storage in electronic databases other than as specified above is not permitted without prior permission in writing from the IUCr.

For further information see <http://journals.iucr.org/services/authorrights.html>



An enhanced set of displacement parameter restraints in *CRYSTALS*

Pascal Parois,* James Arnold and Richard Cooper

Chemical Crystallography, Chemistry Research Laboratory, Mansfield Road, Oxford OX1 3TA, UK. *Correspondence e-mail: pascal@parois.net

Received 28 November 2017

Accepted 10 May 2018

Edited by A. R. Pearson, Universität Hamburg, Germany

Keywords: TLS (translation, libration, screw); disorder; atomic displacement; refinement; crystallographic software; restraints; *CRYSTALS*.

Supporting information: this article has supporting information at journals.iucr.org/j

Crystallographic restraints are widely used during refinement of small-molecule and macromolecular crystal structures. They can be especially useful for introducing additional observations and information into structure refinements against low-quality or low-resolution data (*e.g.* data obtained at high pressure) or to retain physically meaningful parameter values in disordered or unstable refinements. However, despite the fact that the anisotropic displacement parameters (ADPs) often constitute more than half of the total model parameters determined in a structure analysis, there are relatively few useful restraints for them, examples being Hirshfeld rigid-bond restraints, direct equivalence of parameters and *SHELXL* RIGU-type restraints. Conversely, geometric parameters can be subject to a multitude of restraints (*e.g.* absolute or relative distance, angle, planarity, chiral volume, and geometric similarity). This article presents a series of new ADP restraints implemented in *CRYSTALS* [Parois, Cooper & Thompson (2015), *Chem. Cent. J.* **9**, 30] to give more control over ADPs by restraining, in a variety of ways, the directions and magnitudes of the principal axes of the ellipsoids in locally defined coordinate systems. The use of these new ADPs results in more realistic models, as well as a better user experience, through restraints that are more efficient and faster to set up. The use of these restraints is recommended to preserve physically meaningful relationships between displacement parameters in a structural model for rigid bodies, rotationally disordered groups and low-completeness data.

1. Notation¹

\mathbf{U}^* : second-rank tensor of a dimensionless anisotropic displacement parameter (ADP)

\mathbf{U} : 3×3 matrix of an ADP as defined in a CIF, in \AA^2

\mathbf{N} : transformation matrix converting the unitless ADP \mathbf{U}^* into \mathbf{U}

U_{ij} or $[\mathbf{U}]_{ij}$: element (i, j) of the matrix \mathbf{U}

\mathbf{U}_k : ADP of the k th atom

$U_{ij,k}$: element (i, j) of the ADP \mathbf{U}_k of the k th atom

$[\mathbf{U}]_{\mathcal{L}}$: ADP \mathbf{U} expressed in a new basis \mathcal{L}

$[\mathbf{U}]_{\mathcal{C}}$: ADP \mathbf{U} in a Cartesian coordinate system

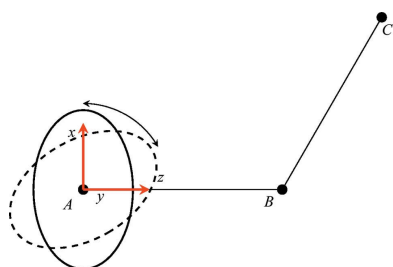
$[\mathbf{U}]_{ij,\mathcal{L}}$: element (i, j) of the ADP \mathbf{U} expressed in a new basis \mathcal{L}

\mathbf{A} : orthogonalization matrix of the atomic coordinates

s.u.: standard uncertainty

2. Introduction

Restraints were first suggested by Waser in 1963. They were initially applied to atomic coordinates and distances, and then extended to isotropic and anisotropic displacement parameters. The displacement parameters of neighbouring atoms can be restrained to agree with values from a rigid-body motion analysis (Rollett, 1970). Since then, very few additional restraints have been implemented, the key development



being the Hirshfeld rigid bond (Hirshfeld, 1976), direct equivalence of parameters and, more recently, *SHELXL* RIGU-type restraints (Thorn *et al.*, 2012). These anisotropic displacement parameter (ADP) restraints in current crystallographic packages are still rather limited. Typically, either two ADPs can be restrained to be equal but without taking into account any rotation between them, or the components of two ADPs along a bond can be restrained to be equal. No general restraint is available for specific relationships between ADPs. ADPs can also be restrained loosely to an isotropic equivalent value in an attempt to avoid physically unreasonable parameter values, but on a strongly anisotropic displacement parameter this restraint is inadequate.

Mean-square displacements of individual atoms can be related to the overall displacement of a rigid group of atoms using the TLS (translation, libration, screw) model (Schomaker & Trueblood, 1968; Dunitz *et al.*, 1988). The TLS model contains 20 independent parameters describing the motion of the rigid group, and is usually fitted to a conventional crystallographic model in order to interpret or validate an existing set of independently refined ADPs. Displacement parameter values predicted by a TLS model can be used as restraints in a crystallographic model, resulting in more physically reasonable displacement parameters, and TLS parameters can even be refined directly against crystallographic data, an approach that has been used in the refinement of resolution-limited macromolecular structures due to the usually poor observable data-to-parameter ratio (Holbrook & Kim, 1984; Howlin *et al.*, 1989; Winn *et al.*, 2001). There are some limitations to this approach: firstly, a TLS model must replace the displacement parameters of at least four atoms in order to provide a simpler description of the displacements using fewer parameters; and secondly, the TLS fit calculation will fail if all the atoms lie in a circle or on any conic section (Cruickshank, 1956; Schomaker & Trueblood, 1968), which rules out simple groups such as CF₃ (one can add the next-neighbour atom to make the TLS fit possible; see §4.1) and larger symmetric groups such as phenyl. However, the program *CRYSTALS* (Parois *et al.*, 2015) has a mechanism to handle the instability in some cases using eigendecomposition, and in practice a TLS fit on a phenyl ring is often possible.

In the first section we concisely describe the mathematical model used and its implementation for each of the ten new restraints, which demonstrate a variety of ways to meaningfully link the components and properties of ADPs in a crystal structure. This serves to document the implementation and also provides a general approach to the design of ADP restraints for future developments. Then, a selection of examples will describe the use of these new restraints in a real case scenario.

3. Implementation of displacement parameter restraints

We designed ten new restraints which can be separated into two groups. (i) Restraints on the direction and/or magnitude

of the principal axes of a single atom relative to each other or to a local coordinate system: UPERP, UPLANE, UQISO, UEIG. (ii) Restraints on the direction and/or magnitude of the principal axes to be similar to, or the average of, other atoms, optionally in a local coordinate system: ULIJ, URIGU, UALIGN, UTLS, UVOL, UEQIV. These new restraints are implemented in *CRYSTALS* (Parois *et al.*, 2015; Version 14.6720 and above) and full details are provided here so that they may be re-implemented in other tools as required.

3.1. Anisotropic displacement parameters

The atomic displacements that define the probability density functions of anisotropic displacements are commonly parameterized as a trivariate Gaussian. One notes that, in the Gaussian approximation, the effect of the mean-square anisotropic displacements (composed of thermal and static contributions) in reciprocal space is fully described in equation (1) by a dimensionless symmetric tensor of rank 2: \mathbf{U}^* or β with $\beta = 2\pi^2\mathbf{U}^*$.

$$\begin{aligned} T(\mathbf{h}) &= \exp(-2\pi^2\mathbf{h}^t\mathbf{U}^*\mathbf{h}), \\ T(\mathbf{h}) &= \exp(-\mathbf{h}^t\beta\mathbf{h}). \end{aligned} \quad (1)$$

By convention, in the output of most crystallographic programs, including the CIF, the six parameters of the ADP have the dimension of [length]² (Grosse-Kunstleve & Adams, 2002) and are obtained using equations (2) and (3) below. \mathbf{U} [denoted \mathbf{U}_{cif} by Grosse-Kunstleve & Adams (2002)] is also the representation stored internally in *CRYSTALS*.

$$\begin{aligned} \mathbf{U} &= \mathbf{N}^{-1}\mathbf{U}^*(\mathbf{N}^{-1})^t, \\ \mathbf{U}^* &= \mathbf{N}\mathbf{U}\mathbf{N}^t, \end{aligned} \quad (2)$$

$$\mathbf{N} = \begin{bmatrix} a^* & 0 & 0 \\ 0 & b^* & 0 \\ 0 & 0 & c^* \end{bmatrix}. \quad (3)$$

3.2. Definition of a Cartesian local coordinate system

A local coordinate system \mathcal{L} is defined for an ADP such that the basis vectors are related to specific directions. In this publication the new coordinate system can be

- (i) a local Cartesian space relative to neighbouring bonds, or
- (ii) a local Cartesian space derived from an averaged orientation of the principal axes of several ADPs.

The 3×3 matrix \mathbf{U} defining the ADP is written in the newly defined coordinate system \mathcal{L} using the transformation matrix \mathbf{T} : $[\mathbf{U}]_{\mathcal{L}} = \mathbf{T}\mathbf{U}\mathbf{T}^t$.

The transformation matrix \mathbf{T} can be decomposed into three parts: the conversion to the dimensionless \mathbf{U}^* tensor using the scaling matrix \mathbf{N} [equation (3)], the orthogonalization matrix \mathbf{A} (having the ADPs expressed in a Cartesian basis simplifies the calculations and removes the need for the metric tensor in geometric transformations), and finally a rotation matrix \mathbf{R} that rotates the basis into the desired position. The transformation matrix \mathbf{T} is therefore $\mathbf{T} = \mathbf{R}\mathbf{A}\mathbf{N}$.

3.3. Derivation of the restraints

Restraints as suggested by Waser (1963) are new observations to be used in addition to the usual X-ray observations. The linearized equation of the restraint G of the parameters x used in least squares is (Watkin, 2008)

$$G_{\text{obs}} - G_{\text{calc}} = \sum_{\text{parameters}} \frac{\partial G_{\text{calc}}}{\partial x} \delta x. \quad (4)$$

If we wish to restrain the values of the ADP in a local basis \mathcal{L} , then the derivatives of the ADP \mathbf{U} need to be expressed as a function of the underlying model parameters U_{ij} . When the atom is lying on a general position, the six partial derivatives $\partial[\mathbf{U}]_{\mathcal{L}}/\partial U_{ij}$ can be derived from the six partial derivatives $\partial\mathbf{U}/\partial U_{ij}$ [equation (5)] using the transformation matrix \mathbf{T} [equation (6)]:

$$\begin{aligned} \frac{\partial\mathbf{U}}{\partial U_{11}} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, & \frac{\partial\mathbf{U}}{\partial U_{22}} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ \frac{\partial\mathbf{U}}{\partial U_{33}} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, & \frac{\partial\mathbf{U}}{\partial U_{12}} &= \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ \frac{\partial\mathbf{U}}{\partial U_{23}} &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, & \frac{\partial\mathbf{U}}{\partial U_{13}} &= \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \end{aligned} \quad (5)$$

$$\frac{\partial[\mathbf{U}]_{\mathcal{L}}}{\partial U_{ij}} = \mathbf{T} \frac{\partial\mathbf{U}}{\partial U_{ij}} \mathbf{T}^t. \quad (6)$$

In *CRYSTALS*, for numerical convenience and speed, the six derived ADPs $\partial\mathbf{U}/\partial U_{ij}$ (3×3 matrix) are vectorized as nine element vectors (Parois & Lutz, 2011). The resulting six vectors (one for each partial derivative) are stacked to form a 9×6 matrix $\partial\mathbf{U}$:

$$\partial\mathbf{U} = \begin{bmatrix} \frac{\partial\mathbf{U}}{\partial U_{11}} & \frac{\partial\mathbf{U}}{\partial U_{22}} & \frac{\partial\mathbf{U}}{\partial U_{33}} & \frac{\partial\mathbf{U}}{\partial U_{12}} & \frac{\partial\mathbf{U}}{\partial U_{13}} & \frac{\partial\mathbf{U}}{\partial U_{23}} \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (7)$$

Each column of $\partial\mathbf{U}$ is a partial derivative $\partial\mathbf{U}/\partial U_{ij}$ [equation (7)]. The six transformations in equation (6) for the six partial derivatives can be rewritten as one expression: $\partial[\mathbf{U}]_{\mathcal{L}} = (\mathbf{T} \otimes \mathbf{T})\partial\mathbf{U}$, where \otimes is the Kronecker product (Parois & Lutz, 2011).

For an atom on a special position, the ADPs will have fewer than six parameters and the derivatives from equation (5) will be different. However, the constraints on U_{ij} from the special positions can also be dealt with at a later stage. In *CRYSTALS*,

TALS, the method above is also used for special positions, and the six partial derivatives as calculated for a general position will be combined using the matrix of constraint (Watkin, 2008) in order to get the $n < 6$ U_{ij} parameters.

4. Description of the new restraints

4.1. UPERP

This restraint restrains the direction of two principal axes of the ADP to be perpendicular to a vector. The vector can be anything, but in *CRYSTALS* it is defined using the atomic model, *i.e.* the atoms, either using two atoms as a vector or using the vector defined by the bisector of two vectors formed by three atoms (Fig. 1). A typical use would be a rotating CF_3 group around C—C or a rotating cyclopentadienyl ring in the plane around its centroid. The libration causes the ADP to be tangential to the rotation axis, which this restraint and *UPLANE* can enforce.

The local coordinate system \mathcal{L} is defined as follows:

- (i) the z axis is along the bond $A-B$ or the bisector of $A-B$ and $A-C$;
- (ii) the x axis is arbitrary but orthogonal to the z axis; and
- (iii) the y axis completes a direct orthonormal coordinate system.

The restraint itself sets the target for two off-diagonal terms $[U]_{13,\mathcal{L}}$ and $[U]_{23,\mathcal{L}}$ of the ADP to zero. In this way, the direction of one principal axis of the ADP is restrained to be aligned with the z axis, while the other two may take any orthogonal directions in the xy plane. The first atom in the command will be restrained (A). The other atoms (B and C) only define the local basis \mathcal{L} .

The syntax of this restraint is

```
UPERP 0.001 A TO B [TO C]
```

4.2. UPLANE

This restraint restrains the direction of a principal axis of the ADP to be perpendicular to the plane of three atoms (Fig. 2). It is most commonly used in combination with *UPERP*

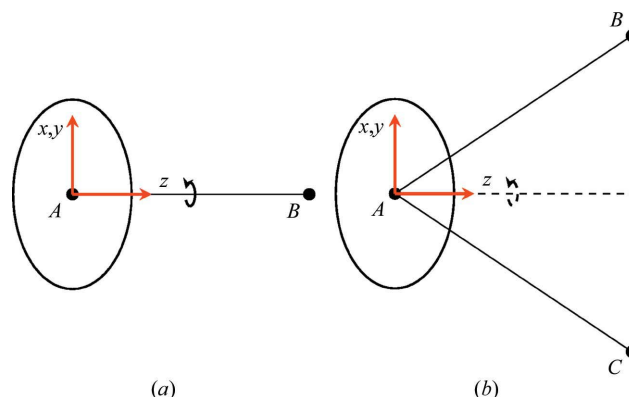


Figure 1
Restraints (a) perpendicular to a bond or (b) perpendicular to the bisector of two bonds (*UPERP*).

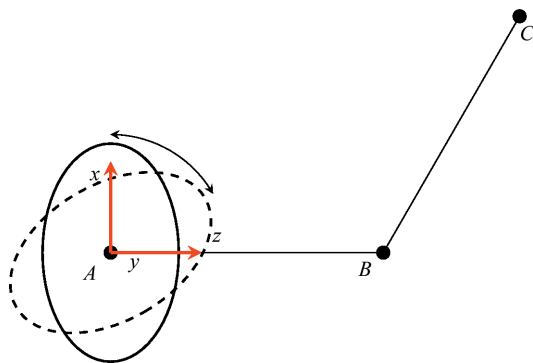


Figure 2

UPLANE restrains the direction of a principal axis of the ADP to be perpendicular to the plane of three atoms.

to enforce the principal axis of an ADP to be tangential to a rotation axis (see §4.1).

The local coordinate system \mathcal{L} is defined as follows:

- (i) the z axis is along the bond $A-B$;
- (ii) the x axis is perpendicular to the z axis and in the plane ABC ; and
- (iii) the y axis completes a direct orthonormal coordinate system.

The restraint itself sets the target for two off-diagonal terms $[U]_{12,\mathcal{L}}$ and $[U]_{23,\mathcal{L}}$ of the ADP. In this way, the direction of one principal axis of the ADP is restrained to be aligned with the y axis, while the other two may take any orthogonal directions in the xz plane. The first atom in the command will be restrained (A). The other atoms (B and C) only define the local basis. Combining the restraints UPERP and UPLANE will result in a fixed orientation of the ADP.

The syntax of this restraint is

UPLANE 0.001 A TO B TO C

4.3. ULIJ

This restraint restrains all the elements of n ADPs A_k to be equal to their average, each in their own local basis \mathcal{L}_k (Fig. 3). The n ADPs will be equal but in different orientations. Each local basis is defined using a further two atoms, B_k and C_k . This restraint is similar to U(IJ) already present in the CRYSTALS software (Parois *et al.*, 2015) and to SIMU present in the SHELXL software (Sheldrick, 2015), which restrain the U_{ij} of two ADPs to be equal, except that this new restraint is applied in a local coordinate system, allowing for more chemically meaningful relationships. This restraint is very useful in cases of high libration and is a further step in combination with UPERP and UPLANE restraints when data are poor.

The local coordinate system of each ADP is defined as follows:

- (i) the z axis is along the bond A_k-B_k ;
- (ii) the x axis is perpendicular to the z axis and in the plane $A_kB_kC_k$; and
- (iii) the y axis completes a direct orthonormal coordinate system.

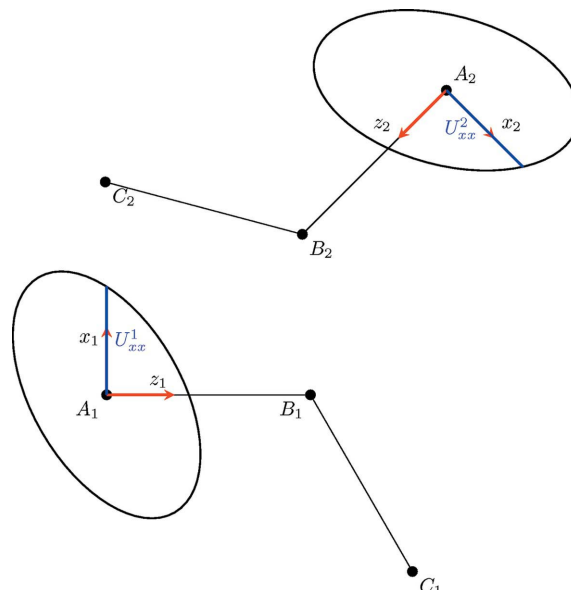


Figure 3

ULIJ restrains all the elements of two ADPs to be equal in their own local coordinate system.

Each ADP \mathbf{U}_k is transformed into a local coordinate system \mathcal{L}_k as defined above. A matrix $\langle[\mathbf{U}]\rangle$ is calculated by averaging the corresponding elements of all the individual $[\mathbf{U}_k]_{\mathcal{L}_k}$ [equation (8)] and then used as the target of $6n$ restraints, one for each element of each $[\mathbf{U}_k]_{\mathcal{L}_k}$ ADP.

$$\langle[\mathbf{U}]\rangle = \frac{1}{n} \sum_{k=1}^n [\mathbf{U}_k]_{\mathcal{L}_k}. \quad (8)$$

The syntax of this restraint is

ULIJ 0.001 A₁ TO B₁ TO C₁, ..., A_n TO B_n TO C_n

4.4. URIGU

These are enhanced rigid-bond restraints implemented as described by Thorn *et al.* (2012).

- (a) $[\mathbf{U}]_{33,\mathcal{L}} = [\mathbf{U}]_{33,\mathcal{L}}$
atom A atom B
- (b) $[\mathbf{U}]_{13,\mathcal{L}} = [\mathbf{U}]_{13,\mathcal{L}}$
atom A atom B
- (c) $[\mathbf{U}]_{23,\mathcal{L}} = [\mathbf{U}]_{23,\mathcal{L}}$
atom A atom B

The local coordinate system is defined as follows:

- (i) the z axis is along the bond $A-B$;
- (ii) the x axis is arbitrary but perpendicular to the z axis; and
- (iii) the y axis completes a direct orthonormal coordinate system.

The syntax of this restraint is

URIGU 0.001 A TO B

4.5. UALIGN

This restraint restrains the direction of all principal axes of n ADPs to be the same but leaves the eigenvalues (*i.e.* the

lengths of the principal axes) free to change. This restraint is an attempt to address disorder in chains, and it is an alternative to the RIGU restraint in *SHELXL* (Thorn *et al.*, 2012). It aligns the principal axes of the ADPs, regardless of differences in the lengths of the principal axes.

Let $[\mathbf{U}]_C = (\mathbf{A}\mathbf{N})\mathbf{U}(\mathbf{A}\mathbf{N})^t$ be the 3×3 matrix of the ADP in an arbitrary Cartesian space. Let \mathbf{Q} be the eigenvectors of $[\mathbf{U}]_C$ and the diagonal elements Λ its eigenvalues. Geometrically, Λ [equation (9)] is the tensor \mathbf{U} in the basis \mathcal{Q} defined by \mathbf{Q} , and the directions of the principal axes of the ellipsoid are aligned with the basis \mathcal{Q} . Because \mathbf{U} is symmetric, \mathbf{Q} is orthogonal and $\mathbf{Q}^{-1} = \mathbf{Q}^t$.

$$\begin{aligned} [\mathbf{U}]_C &= \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1}, \\ \mathbf{\Lambda} &= \mathbf{Q}^{-1}[\mathbf{U}]_C\mathbf{Q}, \\ \mathbf{\Lambda} &= \mathbf{Q}^t[\mathbf{U}]_C\mathbf{Q}, \\ \mathbf{\Lambda} &= (\mathbf{Q}^t\mathbf{A}\mathbf{N})\mathbf{U}(\mathbf{N}^t\mathbf{A}^t\mathbf{Q}). \end{aligned} \quad (9)$$

Let $[\mathbf{U}_k]_C$ be an ADP of the k th atom. Using equation (9), Λ_k are the eigenvalues of $[\mathbf{U}_k]_C$ and \mathbf{Q}_k its eigenvectors. Let $\langle \mathbf{Q} \rangle$ be the average of the n eigenvectors \mathbf{Q}_k and \mathcal{Q}_{avg} the corresponding basis. Let \mathbf{q}_k be the quaternion representation of each eigenvector \mathbf{Q}_k (Shoemaker, 1985). After ensuring that all quaternions lie in the same hemisphere, the average quaternion $\langle \mathbf{q} \rangle = 1/n \sum_{k=1}^n \mathbf{q}_k$ is calculated. Finally, the resulting quaternion is normalized and transformed back to a rotation matrix in order to give $\langle \mathbf{Q} \rangle$.

Each ADP \mathbf{U}_k is then rewritten in the basis \mathcal{Q}_{avg} using $[\mathbf{U}_k]_{\mathcal{Q}_{\text{avg}}} = \mathbf{T}\mathbf{U}_k\mathbf{T}^t$ with $\mathbf{T} = \langle \mathbf{Q} \rangle^t\mathbf{A}\mathbf{N}$. To align the direction of the principal axes of the n ADPs \mathbf{U}_k , it is only necessary to set the target value for off-diagonal terms of $[\mathbf{U}_k]_{\mathcal{Q}_{\text{avg}}}$ to zero using the derivatives from equation (10):

$$\frac{\partial [\mathbf{U}_k]_{\mathcal{Q}_{\text{avg}}}}{\partial U_{ij,k}} = \mathbf{T} \frac{\partial \mathbf{U}_k}{\partial U_{ij,k}} \mathbf{T}^t. \quad (10)$$

The syntax of this restraint is

$$\text{UALIGN } 0.001 \ A_1 \dots A_n$$

4.6. UVOL and UEQIV

The two restraints UVOL and UEQIV restrain the geometric mean and arithmetic mean, respectively, of the magnitude of the principal axes of the ADPs of n atoms to be equal. The geometric mean is proportional to the volume of the representative ellipsoid (hence the name UVOL) and in *CRYSTALS* U_{eq} is defined using the arithmetic mean. ADPs of identical neighbouring atom types usually have a similar mean. It is a very soft restraint, as it does not impose any orientation or shape, but it can help deal with non-positive-definite ADPs.

Let \mathbf{Q} be the eigenvectors of $[\mathbf{U}]_C$ and Λ its eigenvalues, as in equation (9). The geometric mean is $\langle \Lambda \rangle_g = (\Lambda_{11}\Lambda_{22}\Lambda_{33})^{1/3}$ and the arithmetic mean is $\langle \Lambda \rangle_a = \frac{1}{3}(\Lambda_{11} + \Lambda_{22} + \Lambda_{33})$.

The partial derivatives of the geometric mean as a function of U_{ij} are given in equation (11) using equations (6) and (10)

with $\mathbf{T} = \mathbf{Q}^t\mathbf{A}\mathbf{N}$, and $[\mathbf{T}(\partial\mathbf{U}/\partial U_{ij})\mathbf{T}^t]_{ij}$ is the (i, j) element of the tensor $\mathbf{T}(\partial\mathbf{U}/\partial U_{ij})\mathbf{T}^t$.²

$$\begin{aligned} \frac{\partial \langle \Lambda \rangle_g}{\partial U_{ij}} &= \frac{1}{3(\langle \Lambda \rangle_g)^2} \left(\Lambda_{22}\Lambda_{33} \left[\mathbf{T} \frac{\partial \mathbf{U}}{\partial U_{ij}} \mathbf{T}^t \right]_{11} \right. \\ &\quad + \Lambda_{11}\Lambda_{33} \left[\mathbf{T} \frac{\partial \mathbf{U}}{\partial U_{ij}} \mathbf{T}^t \right]_{22} \\ &\quad \left. + \Lambda_{11}\Lambda_{22} \left[\mathbf{T} \frac{\partial \mathbf{U}}{\partial U_{ij}} \mathbf{T}^t \right]_{33} \right). \end{aligned} \quad (11)$$

Similarly, the same expression is derived for the arithmetic mean in equation (12),

$$\frac{\partial \langle \Lambda \rangle_a}{\partial U_{ij}} = \frac{1}{3} \left(\left[\mathbf{T} \frac{\partial \mathbf{U}}{\partial U_{ij}} \mathbf{T}^t \right]_{11} + \left[\mathbf{T} \frac{\partial \mathbf{U}}{\partial U_{ij}} \mathbf{T}^t \right]_{22} + \left[\mathbf{T} \frac{\partial \mathbf{U}}{\partial U_{ij}} \mathbf{T}^t \right]_{33} \right). \quad (12)$$

The syntax of these restraints is

$$\begin{aligned} \text{UVOL } 0.001 \ A_1 \dots A_n \\ \text{UEQIV } 0.001 \ A_1 \dots A_n \end{aligned}$$

4.7. UEIG

The two principal axes of an ADP with the smallest magnitudes are restrained to be equal. This restraint is also designed to deal with non-positive-definite ADPs, so it is a more flexible approach than UQISO (which restrains all principal axes to be equal; see §4.8) and is suitable where strong anisotropy is observed.

Let \mathbf{Q} be the eigenvectors of $[\mathbf{U}]_C$ and the diagonal elements of Λ its eigenvalues, as in equation (9).

The eigendecomposition is sorted so that the eigenvalues are in ascending order. The two values Λ_{11} and Λ_{22} are restrained to be equal and the cross term Λ_{12} is restrained to be zero to ensure two-dimensional isotropic distribution.

The syntax of this restraint is

$$\text{UEIG } 0.001 \ A$$

4.8. UQISO

All magnitudes of the principal axes of an ADP are restrained to be equal. With a high weight (or low estimated error) this restraint will make an ADP close to isotropic, so it is similar to the ISOR restraint in *SHELXL* (Sheldrick, 2015).

Let \mathbf{Q} be the eigenvectors of $[\mathbf{U}]_C$ and the diagonal elements of Λ its eigenvalues, as in equation (9).

Under the restraint, Λ_{11} , Λ_{22} and Λ_{33} are restrained to their mean, and off-diagonal terms of Λ are restrained to zero to ensure sphericity. Six restraints will be generated.

The syntax of this restraint is

$$\text{UQISO } 0.01 \ A$$

² Reminder: let $z = (uvw)^{1/3}$. Then $z' = (vuw' + uvw' + uvw')/[3(uvw)^{2/3}]$ with $u = [\Lambda(U_{ij})]_{11}$, $v = [\Lambda(U_{ij})]_{22}$ and $w = [\Lambda(U_{ij})]_{33}$, where the prime (') denotes differentiation.

4.9. UTLS

The restraint UTLS assumes a group of atoms to be a rigid body and fits a TLS model against the current parameters of n atoms. Given the TLS model, a new set of ADPs is calculated for the first m atoms and used as a target for the restraint. If m is omitted when writing the restraint, all the atoms used in calculating the model will be restrained.

Note that, as the target is computed from a new model in each cycle, this will slow down convergence dramatically.

The syntax of this restraint is

```
UTLS 0.01 [m] A1 ... Am ... An
```

4.10. Choosing s.u. values for the restraints

Unfortunately, there is no method to choose an s.u. for a restraint automatically and any guidance value may be too controversial. There is, however, a way of measuring the strength of a restraint using leverage. The command #CHECK in CRYSTALS will report the leverages of every restraint (P. Parois & R. Cooper, manuscript in preparation). A value close to zero means that the restraint has no effect on the model and a leverage value close to one means that the restraint fully determines the parameters (*i.e.* the X-ray data have no influence) (Parsons *et al.*, 2012). The user can then adjust the s.u. of a restraint to the desired effect: a very high leverage for very poor data or a more subtle restraint to guide the existing data.

5. Examples

5.1. Disordered CF₃

Disordered trifluoromethyl groups are very common and, if the minor component has a very low occupancy, the displacement parameters become ill defined. The crystal structure of Ar''₂PBr, denoted **4** in the work of Batsanov *et al.* (2002), shows a disordered CF₃ group on atom C141. The major part, atoms F7A, F8A and F9A, has a fixed occupancy of 0.90, while the minor component is composed of atoms F7B, F8B and F9B. Atoms in opposite (major and minor) positions were refined with equivalent ADPs. This can lead to two potential

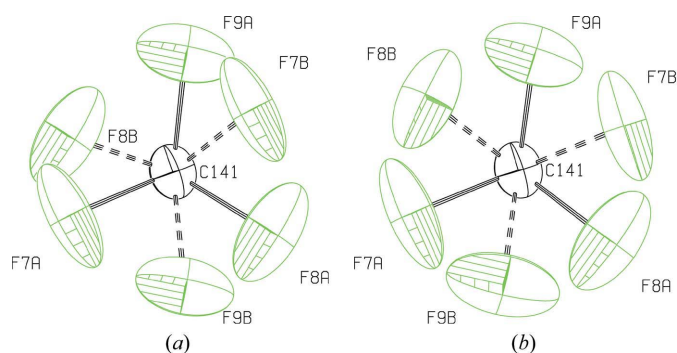


Figure 4
Comparison with (a) equivalent ADPs and (b) the UTLS restraint. Naming has been changed to match the original publication using F7A → F71, F8A → F81, F9A → F91, F7B → F72, F8B → F82 and F9B → F92.

problems: (i) if the two fluorine atoms are not exactly 180° apart, the ADP is no longer tangential to the rotation; and (ii) the ADP is not mirrored through the symmetry. An isotropic model for the minor component could be used but is unsatisfactory because of the high libration around the C—C bond, which makes the ADPs very anisotropic. Because of technical limitations in CRYSTALS, the atom numbering used by Batsanov *et al.* (2002) could not be kept and has been changed as follows: F7A → F71, F8A → F81, F9A → F91, F7B → F72, F8B → F82 and F9B → F92.

Restraining the atomic coordinates so that the opposite positions are perfectly mirrored is possible using distance restraints and is a reasonable assumption. Restraints could be used on ADPs at opposite positions to impose symmetry on the ADPs. However, this configuration assumes that the libration axis is on the C—C bond, which is not the case here: TLS analysis finds a libration axis at an angle of 6.6° from the C—C bond.

Instead, the minor component can be restrained using a more general restraint based on the TLS restraint: UTLS. The TLS model is calculated on all the atoms involved (C14, C141, F71, F81, F91, F72, F82 and F92), but only applied to the fluorine atoms from the minor component (F72, F82 and F92).

A fairly high weight is necessary, because of the very low occupancy of the minor component. In this case, the restraint behaves very closely to a constraint to a calculated ADP. Also, because both the major and minor components are used in the same TLS model, the assumption is that the minor component will have the same libration as the major component. Distance restraints were used so that the two components are symmetrical (this is only necessary to avoid the merging of the two components), and the occupancy refined to a value of 0.919 (2). The full list of restraints is available in the supporting information.

The syntax of this restraint is

```
UTLS 0.001 3 F72 F82 F92
CONT C14 C141 F71 F82 F91
```

The model of the trifluoromethyl group from the work of Batsanov *et al.* (2002) using EADP constraints is compared with

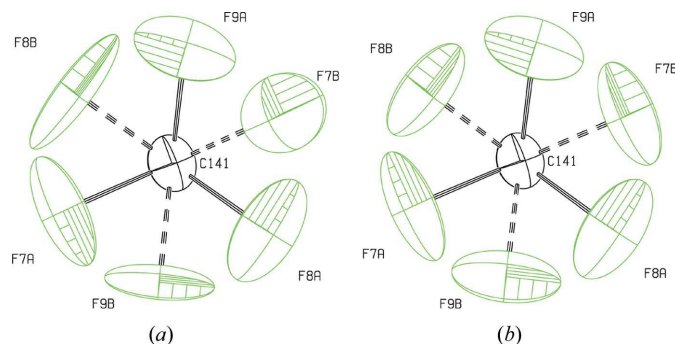


Figure 5
Comparison with (a) UPERP and UPLANE restraints, and (b) UPERP, UPLANE and ULIJ restraints. Naming has been changed to match the original publication using F7A → F71, F8A → F81, F9A → F91, F7B → F72, F8B → F82 and F9B → F92.

the model obtained above in Fig. 4. The ADPs using the UTLS restraint are more physically meaningful.

For comparison, we also performed a combination of UPERP, UPLANE and ULIJ restraints of the minor component (Fig. 5).

The syntax of the restraints above is

```
UPERP 0.001 F72 TO C141
UPERP 0.001 F82 TO C141
UPERP 0.001 F92 TO C141
UPLANE 0.001 F72 TO C141 TO C14
UPLANE 0.001 F82 TO C141 TO C14
UPLANE 0.001 F92 TO C141 TO C14
ULIJ 0.001 F72 TO C141 TO C14
CONT F82 TO C141 TO C14
CONT F92 TO C141 TO C14
```

The restraints UPERP and UPLANE were not sufficient alone to get a meaningful model, and ULIJ had to be used because the longest axis of the ADP on atom F7B is perpendicular to the libration, which is very unlikely. The full list of restraints is available in the supporting information.

Refinement using UTLS restraints, or with a combination of UPERP, UPLANE and ULIJ restraints, gives a similar R factor of 0.035 and an occupancy of 0.08 for the minor component. They both also give a good model of the disorder. The UTLS restraint allows more freedom for the orientation of the ADPs, which is beneficial if the rotation axis deviates significantly from the C—C bond.

A search of the Cambridge Structural Database (CSD) (Groom *et al.*, 2016; Version 5.38 with updates up to February 2017), limited to IUCr journals³ and using the C—CF₃ motif, yielded 343 hits after data processing (see supporting information). Because ADPs are not present in the CSD, the corresponding CIFs were used to fit a model of the ADPs to the deposited ADPs and calculate the R_1 factor [equation (13)] for the U values (Schomaker & Trueblood, 1968) on each hit.

$$R_1 = \frac{\sum_{k=1}^{\text{atoms}} \sum_{\substack{i=1 \\ j=1 \\ i \leq j}}^3 \left| \frac{U_{ij,k}}{\text{atomic model}} - \frac{U_{ij,k}}{\text{TLS model}} \right|}{\sum_{k=1}^{\text{atoms}} \sum_{\substack{i=1 \\ j=1 \\ i \leq j}}^3 \left| \frac{U_{ij,k}}{\text{atomic model}} \right|}. \quad (13)$$

In the search described above, 80% of the structures present an R_1 factor for the U values of below 0.04 (Fig. 6), which means that the UTLS restraint is well suited to the CF₃ group.

5.2. High-pressure refinement

High-pressure data suffer from low completeness because of detector shadowing by the pressure cell, plus higher noise

³ The limit to a single publisher is because ADPs are not stored in the CSD, whereas most articles in IUCr journals have their structures deposited with the IUCr and these are easily obtained *via* a script.

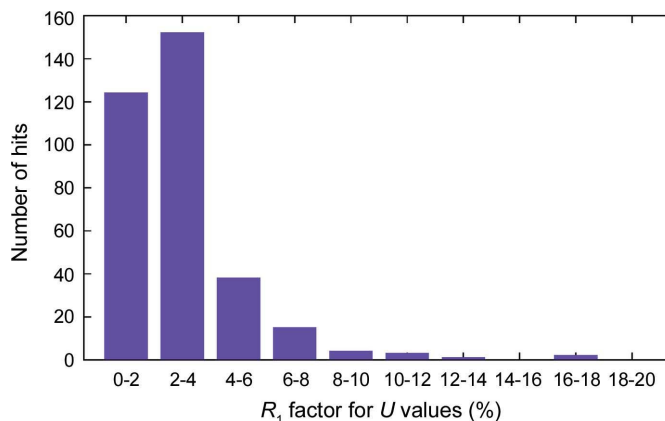


Figure 6
Frequency distribution of R_1 factors from the fit of a TLS model for a C—CF₃ bond in a selection of published structures

and systematic errors due to a higher background and parasitic reflections. Complete or partial isotropic refinement may be the only option during the refinement.

Madsen *et al.* (2016) reported a high-pressure crystallographic study of a coordination polymer, emim[Mn^{II}(btc)] (emim is 1-ethyl-3-methylimidazolium and btc is 1,3,5-benzenetricarboxylate). In particular, in the structure at 3.2 GPa the authors decided to refine the btc ligand isotropically after seeing no improvement with an anisotropic refinement of the ligand. However, between isotropic and anisotropic refinement there are a range of possible intermediate solutions that can be explored using a set of stronger and stronger restraints from an anisotropic refinement.

Combining rigid-body restraints, quasi-isotropic restraints and U_{eq} restraints, it is possible to push the boundary where isotropic refinement becomes necessary. These restraints are general enough to be applied to a whole structure or a single molecule, and they are also very easy and quick to set up. In this example

(i) all the ADPs were restrained to be approximately isotropic using UQISO,

(ii) all the ADPs were restrained to have the same U_{eq} as a neighbouring atom using UEQIV and

(iii) all the ADPs were restrained to have the same contribution along the bond using rigid-body restraints (VIBR).

As a result, when the estimated errors of the UQISO, UEQIV and VIBR restraints are very low (*i.e.* weights are very high), the fit to the X-ray data model will be equivalent to a completely isotropic model (s.u. of the order of 10^{-4}). At the other extreme, when the estimated errors are high the model will be equivalent to an unrestrained anisotropic model (s.u. of the order of 0.25).

The structure was refined using several s.u. values between these two extremes. As expected, the R and wR_2 factors⁴ drop when the restraints are loose as the model has more freedom. However R and wR_2 do not give a warning for overfitting,

⁴ $R = (\sum |F_{obs} - F_{calc}|) / \sum |F_{obs}|$ and $wR_2 = \{[\sum w(F_{obs}^2 - F_{calc}^2)^2] / [\sum w(F_{obs}^2)]\}^{1/2}$, where F_{obs} and F_{calc} are structure-factor magnitudes.

which is very likely in this case owing to the low completeness. The ratio of parameters to observations cannot be used in this context as the restraints are not equivalent to a true observation. In an extreme case, the ADPs become non-positive definite or otherwise physically meaningless, and this can be used as a warning for overfitting. A better estimate of the fit is therefore needed, which takes overfitting into account. A method similar to R_{complete} was used here to check if the restraints did indeed improve the solution (Luebben & Gruene, 2015).

The list of 870 merged reflections was partitioned into ten random parts of 87 reflections each. A working data set was formed using nine parts and the remaining part was used as a test set. Ten sets are formed from the ten possibilities of arranging the parts. Using each working set, the model is refined and wR_2 calculated on the test set only. The ten values of wR_2 were then averaged: $\langle wR_2 \rangle$ and the variance $\sigma(\langle wR_2 \rangle)^2$ were calculated. The variance was fairly high because the test set was a significant number of reflections removed from the refinement. Therefore, we repeated the procedure n times with a different sampling each time and calculated a weighted average using $\sigma(\langle wR_2 \rangle_i)^2$ as the weight for each $\langle wR_2 \rangle_i$.

Finally, the whole procedure above was repeated using different s.u. values for the restraints. A complete model including restraints, with s.u. values ranging from 0.01 to 0.08, was determined to start with. These s.u. values are then multiplied by a scaling factor ranging from 0.01 (equivalent to isotropic, all ADPs the same) to 3.0 (nearly anisotropic) to explore the effect of restraints on the refinement.

Fig. 7 shows that wR_2 improves for the working data set as the model becomes more flexible. However, wR_2 for the test set improves slightly, then gets much worse as the flexibility increases. This minimum in the test set ($wR_2 \simeq 0.27$) is a

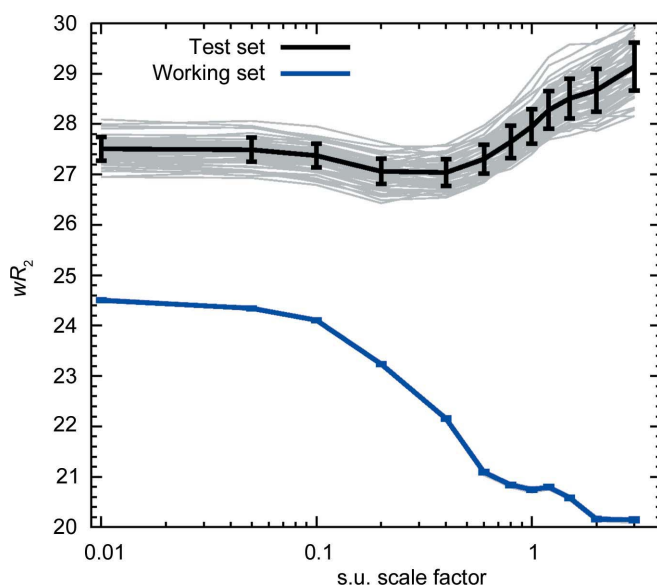


Figure 7

Values of wR_2 of the test set and working set, plotted as a function of the strength of the restraints. The s.u. scale factor is a multiplier applied to our best guess of s.u. values for the restraints.

balance between underfitting (model too simple) and overfitting (model has too much freedom).

5.3. Restraints on disordered pentamethylcyclopentadienyl

When disorder is present the quality of the data often suffers, and the more diffuse electron density makes it more difficult to refine part of the structure. Restraints are very useful in this case to keep the model parameters physically reasonable. The structure WM5234 (Becker *et al.*, 2015) has two pentamethylcyclopentadienyl ligands, including one disordered over two sites with refined occupancies of 0.811 (3):0.189 (3). The authors refined the disordered ring using distance and angle restraints, plus equality constraints (EADP) on the ADPs of atoms C17A–C21A, C17B–C21B, and C22A–C26A, C22B–C26B (see Fig. 8).

The pentamethylcyclopentadienyl molecule is rigid and must behave like a rigid group. Therefore, the orientation and relative size of the ADPs is very unlikely. A search of the CSD, again limited to IUCr journals, yielded 97 hits in 66 structures (see supporting information). Because ADPs are not present in the CSD, the corresponding CIFs were retrieved to perform a TLS analysis (Schomaker & Trueblood, 1968) on each hit.

In this search, 70% of the structures give a fit to the TLS model with an R_1 factor between 0.03 and 0.09 (Fig. 9), and a visual examination of the worst fit revealed some non-meaningful physical parameters. These results confirmed our assumption of a rigid molecule.

We also performed the same analysis on the non-disordered and two disordered pentamethylcyclopentadienyl groups from the structure WM5234 as published and we obtained values of 0.0475, 0.1664 and 0.3530, respectively, for the R factors of the fit of a TLS model. The results show a significant deviation from a rigid body for the two disordered parts.

A restraint on the ADPs based on a rigid body is possible, but the cost of calculating the TLS model at each refinement iteration is significant and the ever-changing target slows down

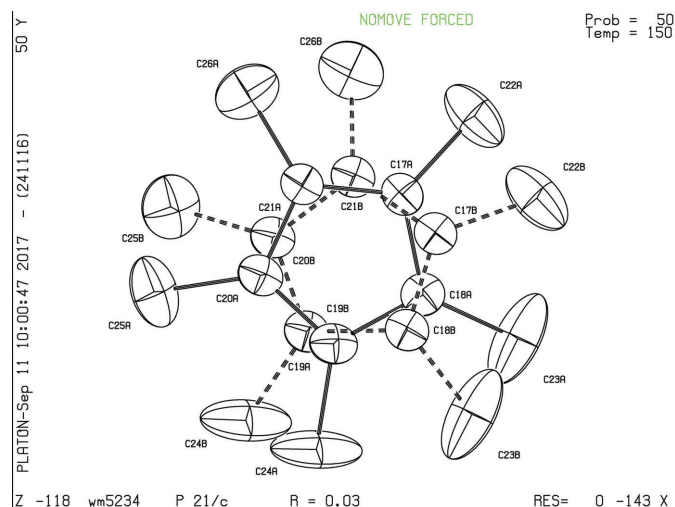


Figure 8

A PLATON plot (Spek, 2009) of the disordered pentamethylcyclopentadienyl ring from the published structure of Becker *et al.* (2015).

convergence significantly. As a very rough estimate, it can require up to ten times as many cycles to converge. In this particular case the situation is even worse, as convergence does not happen for the minor component because of its very low occupancy of below 0.2. It is, however, possible to mimic such a restraint using simpler ADP restraints, assuming the libration axis of the TLS model of the pentamethylcyclopentadienyl ligand to be normal to the plane of the molecule and through its centroid. The libration then implies that the ADPs must be tangential to the axis of rotation, and the ADPs of the same molecule at the same distance to the rotation axis should be identical. This results in 22 restraints per disordered pentamethylcyclopentadienyl ligand: ten UPLANE restraints within the plane of the molecule, ten UPERP restraints perpendicular to the direction of the rotation axis and two ULIJ equality restraints (see supporting information).

After refinement with the new restraints (Fig. 10), a TLS analysis of the pentamethylcyclopentadienyl groups shows that, for the major component (occupancy of 0.802), the *R* factor for the TLS fit is 0.0285, whereas for the minor

component it is 0.0301. This result is a good improvement compared with the original published structure and the ADPs are now compatible with the libration present in the disordered molecule.

The difference in the magnitude of the disorder between the two conformations of the pentamethylcyclopentadienyl ring can be attributed to the slightly different environments of the conformations in the crystal structure. However, given the low occupancy of the minor component, the model could well be fitting some noise, or it may be a consequence of the correlation between the occupancy and the volume of the ADPs. Comparing the longest magnitude of the different ADPs in both components, we note that they are consistent: both the inner ring and the outer ring of the minor component are longer, indicating a genuine difference.

6. Conclusions

More than ten different restraints have been added to *CRYSTALS*. The result gives the crystallographer more control over the restraint of ADPs and avoids the misuse of inappropriate restraints because of a lack of choice, as exemplified by the use of equivalent ADPs relating atoms C23A and C23B in Fig. 8.

The restraint based on a TLS model is the most powerful of the new restraints because of its simplicity of use, its performance on rigid groups and its physical meaning, although slower convergence could be problematic for larger structures. UQISO, UEQIV and UVOL are also very easy to use and can be very useful for low-quality or low-completeness data. They push the limit on the requirement of an isotropic refinement for low-completeness data sets, *e.g.* high-pressure or excited-state experiments, and can reveal anisotropic features in the crystal structure.

The rest of the restraints are designed to fine tune ADPs as the crystallographer can restrain the orientation and size of any ADP. They do, however, require more consideration to set up and will require appropriate justification if used in a refinement.

As noted in §5.2, we should be striving to construct the model which best represents the real structure, rather than necessarily that which gives the best *R* factor.

These new restraints are available in *CRYSTALS* Version 14.6720 and above, available on the Oxford Chemical Crystallography web site, <http://www.xtl.ox.ac.uk>.

Funding information

PP and RC acknowledge funding from the EPSRC (grant No. EP/K013009/1).

References

- Batsanov, A. S., Cornet, S. M., Dillon, K. B., Goeta, A. E., Hazendonk, P. & Thompson, A. L. (2002). *J. Chem. Soc. Dalton Trans.* pp. 4622–4628.
- Becker, L., Spannenberg, A., Arndt, P. & Rosenthal, U. (2015). *Acta Cryst. E71*, m219–m220.

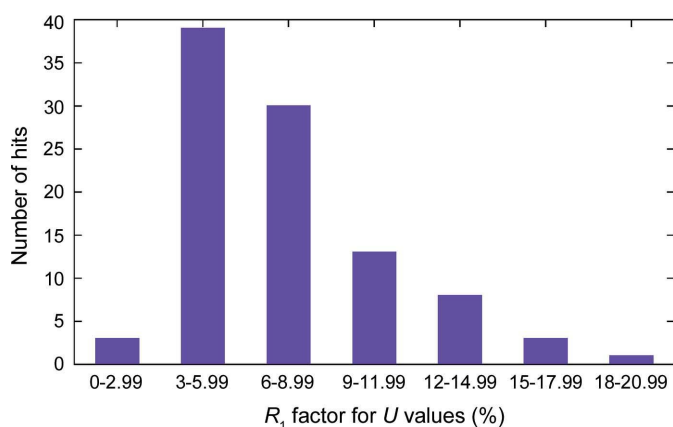


Figure 9
Frequency distribution of *R*₁ factors from the fit of a TLS model for a pentamethylcyclopentadienyl ring in a selection of published structures.

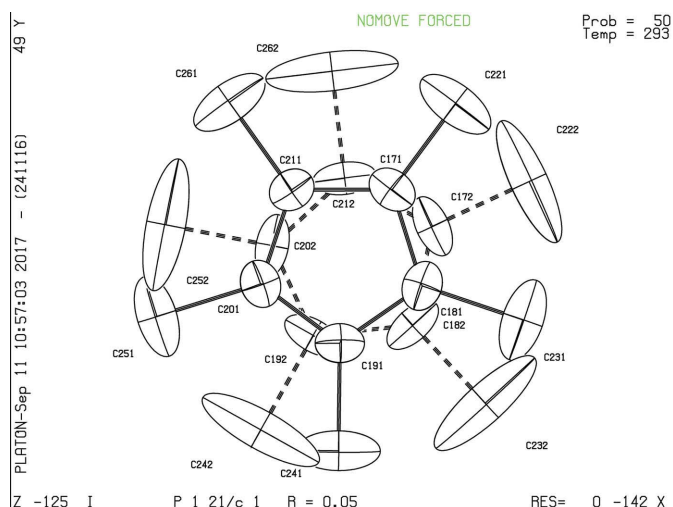


Figure 10
A *PLATON* plot of the disordered pentamethylcyclopentadienyl ring after refinement as a rigid body.

- Cruickshank, D. W. J. (1956). *Acta Cryst.* **9**, 754–756.
- Dunitz, J. D., Schomaker, V. & Trueblood, K. N. (1988). *J. Phys. Chem.* **92**, 856–867.
- Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. (2016). *Acta Cryst.* **B72**, 171–179.
- Grosse-Kunstleve, R. W. & Adams, P. D. (2002). *J. Appl. Cryst.* **35**, 477–480.
- Hirshfeld, F. L. (1976). *Acta Cryst.* **A32**, 239–244.
- Holbrook, S. R. & Kim, S.-H. (1984). *J. Mol. Biol.* **173**, 361–388.
- Howlin, B., Moss, D. S. & Harris, G. W. (1989). *Acta Cryst.* **A45**, 851–861.
- Luebben, J. & Gruene, T. (2015). *Proc. Natl Acad. Sci. USA*, **112**, 8999–9003.
- Madsen, S. R., Moggach, S. A., Overgaard, J. & Brummerstedt Iversen, B. (2016). *Acta Cryst.* **B72**, 389–394.
- Parois, P., Cooper, R. I. & Thompson, A. L. (2015). *Chem. Cent. J.* **9**, 30.
- Parois, P. & Lutz, M. (2011). *Acta Cryst.* **A67**, 383–390.
- Parsons, S., Wagner, T., Presly, O., Wood, P. A. & Cooper, R. I. (2012). *J. Appl. Cryst.* **45**, 417–429.
- Rollett, J. S. (1970). *Crystallographic Computing*, edited by F. R. Ahmed, S. R. Hall & C. P. Huber, pp. 167–181. Copenhagen: Munksgaard.
- Schomaker, V. & Trueblood, K. N. (1968). *Acta Cryst.* **B24**, 63–76.
- Sheldrick, G. M. (2015). *Acta Cryst.* **C71**, 3–8.
- Shoemaker, K. (1985). *Comput. Graph.* **19**, 245–254.
- Spek, A. L. (2009). *Acta Cryst.* **D65**, 148–155.
- Thorn, A., Dittrich, B. & Sheldrick, G. M. (2012). *Acta Cryst.* **A68**, 448–451.
- Waser, J. (1963). *Acta Cryst.* **16**, 1091–1094.
- Watkin, D. (2008). *J. Appl. Cryst.* **41**, 491–522.
- Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* **D57**, 122–133.