From deep TLS validation to ensembles of atomic models built from elemental motions. Addenda and corrigendum

Alexandre Urzhumtsev,[a,b]* Pavel V. Afonine,[c] Andrew H. Van Benschoten,[d] James S. Fraser[d] and Paul D. Adams[c,e]

[a]Centre for Integrative Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire, 1 rue Laurent Fries, BP 10142, 67404 Illkirch, France, [b]Faculté des Sciences et Technologies, Université de Lorraine, BP 239, 54506 Vandoeuvre-les-Nancy, France, [c]Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA, [d]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, California, USA, and [e]Department of Bioengineering, University of California, Berkeley, Berkeley, California, USA. *Correspondence e-mail: sacha@igbmc.fr

Researcher feedback has indicated that in Urzhumtsev *et al.* [(2015) *Acta Cryst.* D**71**, 1668–1683] clarification of key parts of the algorithm for interpretation of TLS matrices in terms of elemental atomic motions and corresponding ensembles of atomic models is required. Also, it has been brought to the attention of the authors that the incorrect PDB code was reported for one of test models. These issues are addressed in this article.

1. Incorrect PDB code

In the original article (Urzhumtsev *et al.*, 2015), we used several atomic models in order to test the algorithms and provide examples. Unfortunately, the incorrect PDB code had been reported for one of them. Everywhere in the text (§6.2, Tables 2 and 3), 1rge should be used instead of 1dqv and ribonuclease S should be used instead of synaptotagmin. We apologize for this confusion. The diffraction data set used for test refinement of ribonuclease S was obtained from the *CCP*4 (Winn *et al.*, 2011) distribution (http://www.ccp4.ac.uk/examples/rnase/rnase25.mtz).

2. Origin choice

The problem of the origin choice is discussed in detail in the review of Urzhumtsev *et al.* (2013) leading us to provide less detail in Urzhumtsev *et al.* (2015). As mentioned in §2.2 of Urzhumtsev *et al.* (2015), the **T** and **S** matrices depend on the point (origin of the TLS group) with respect to which the three libration axes are defined. This is also important for generating the U_n matrices from a set of TLS matrices. Confusion arises from the fact that the TLS origin may be, and in fact usually is, different from the origin of the coordinate system in which the atomic coordinates are provided.

The choice of the TLS origin is arbitrary; some typical choices are described in §2.2. Let $(\tilde{x}_n, \tilde{y}_n, \tilde{z}_n)$, $n = 1, \ldots N$ be atomic Cartesian coordinates as the input parameters of the procedure; for example, they may be the coordinates given in the PDB file. Let $(x_{\mathrm{TLS}}, y_{\mathrm{TLS}}, z_{\mathrm{TLS}})$ be respective coordinates of the origin of the TLS group. The origin of basis [M] defined in Urzhumtsev *et al.* (2015) is assumed to coincide with this point. This means that the coordinates that are input to the ensemble generating procedure ($\mathbf{r}_{[M]n}$ in §7.2 of Urzhumtsev *et*

mean relative difference in U matrices

number of generated models
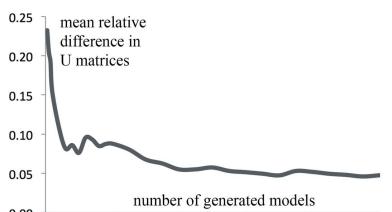
al., 2015) are also expected to be shifted to the origin of the TLS group as follows

$$\left(x_{[M]n}, y_{[M]n}, z_{[M]n}\right) = (\tilde{x}_n, \tilde{y}_n, \tilde{z}_n) - (x_{TLS}, y_{TLS}, z_{TLS}), n = 1, \dots N. \quad (1)$$

The matrix

$$A_n = \begin{pmatrix} 0 & z_{[M]n} & -y_{[M]n} \\ -z_{[M]n} & 0 & x_{[M]n} \\ y_{[M]n} & -x_{[M]n} & 0 \end{pmatrix} \quad (2)$$

[equation (3) in Urzhumtsev *et al.*, 2015] uses these new coordinates (1).

3. U_n matrices

The TLS model is valid for harmonic motions and, as a consequence, for small libration amplitudes only. It allows for calculation of the individual atomic displacement parameters U_n in two different ways. They may be calculated analytically using the formulae (2) and (3) from Urzhumtsev *et al.* (2015). Alternatively, the same matrices can be calculated numerically from the coordinates of the set of models generated explicitly using the procedure described in §7 and *Appendix A* of Urzhumtsev *et al.* (2015). We stress that in formulae (59)–(61) the expressions

$$\left(\Delta_{[L]}^{lx}x; \Delta_{[L]}^{lx}y; \Delta_{[L]}^{lx}z\right) = \{s_x d_{x0}; \left[(y_{[L]} - w_y^{lx})(\cos d_{x0} - 1)\right.$$
$$\left. - (z_{[L]} - w_z^{lx})\sin d_{x0}\right]; \left[(y_{[L]} - w_y^{lx})\right.$$
$$\left. \times \sin d_{x0} + (z_{[L]} - w_z^{lx})(\cos d_{x0} - 1)\right]\}$$

$$\left(\Delta_{[L]}^{ly}x; \Delta_{[L]}^{ly}y; \Delta_{[L]}^{ly}z\right) = \{\left[(z_{[L]} - w_z^{ly})\sin d_{y0} + (x_{[L]} - w_x^{ly})\right.$$
$$\left. \times (\cos d_{y0} - 1)\right]; s_y d_{y0}; \left[(z_{[L]} - w_y^{ly})\right.$$
$$\left. \times (\cos d_{y0} - 1) - (x_{[L]} - w_x^{ly})\sin d_{y0}\right]\}$$

$$\left(\Delta_{[L]}^{lz}x; \Delta_{[L]}^{lz}y; \Delta_{[L]}^{lz}z\right) = \{\left[(x_{[L]} - w_x^{lz})(\cos d_{z0} - 1) - (y_{[L]} - w_y^{lz})\right.$$
$$\left. \times \sin d_{z0}\right]; \left[(x_{[L]} - w_y^{lz})\sin d_{z0}\right.$$
$$\left. + (y_{[L]} - w_y^{lz})(\cos d_{z0} - 1)\right]; s_z d_{z0}\} \quad (3)$$

are the coordinates of the libration shifts in the basis [L]; similarly, the values (t_{x0}, t_{y0}, t_{z0}) in (48) are the coordinates of the vibration shifts in the basis [V]. These coordinates must be converted into the basis [M] in order to obtain the coordinates of the total shifts $(\Delta x_{nk}, \Delta y_{nk}, \Delta z_{nk})$, $n = 1, \dots N$, to be applied to the atomic coordinates $(\tilde{x}_n, \tilde{y}_n, \tilde{z}_n)$, $n = 1, \dots N$. Here k is the number of the generated model.

Once an ensemble is generated, the atomic displacement matrix U_n for each atom n

$$U_n = \begin{pmatrix} \langle \Delta x_n^2 \rangle & \langle \Delta x_n \Delta y_n \rangle & \langle \Delta x_n \Delta z_n \rangle \\ \langle \Delta x_n \Delta y_n \rangle & \langle \Delta y_n^2 \rangle & \langle \Delta y_n \Delta z_n \rangle \\ \langle \Delta x_n \Delta z_n \rangle & \langle \Delta y_n \Delta z_n \rangle & \langle \Delta z_n^2 \rangle \end{pmatrix} \quad (4)$$

can be calculated directly from the coordinates (x_{nk}, y_{nk}, z_{nk}), $k = 1, \dots K$ of the multiple copies of the same atom in the ensemble [see formula (2.2) from Urzhumtsev *et al.*, 2013]. Here atomic coordinates are in the basis [M] and averaging is performed over all K instances of the atom in the ensemble.

These two somewhat independent routes to obtain the U_n matrices allow a convenient way of validating the described procedures. Indeed, given parameters of elemental motions one can construct TLS matrices and then calculate U_n from these TLS matrices using the analytical expression cited above. Also, one can use TLS matrices to generate an ensemble of models and then derive U_n from the ensemble using formula (4). We added this comparison to *cctbx* (Grosse-Kunstleve *et al.*, 2002) as a test-exercise of the implementation.

Now the *Phenix* (Adams *et al.*, 2010) command `phenix.tls_as_xyz model.pdb n_models = N` creates three PDB files containing the following.

(i) An ensemble of N models (N can be any positive integer) that are consistent with the TLS model (TLS records must be present in the model.pdb file header).

(ii) A single model with anisotropic U_n (U_{TLS} in ANISOU records) calculated from the TLS matrices analytically.

(iii) A single model with anisotropic U_n ($U_{ensemble}$, ANISOU records) calculated numerically from the ensemble of models.

U_n obtained using the two different approaches are expected to be similar with possible differences arising from several sources, such as the following.

(*a*) Non-linearity of the TLS approximation (large libration amplitudes; Table 1).

(*b*) A finite number of models in the ensemble; empirically we found that 5000–10 000 is sufficient most of the time (Fig. 1).

(*c*) Numerical errors arising from a long chain of transformations: from decomposing TLS matrices into basic parameters of elemental motions and using these parameters to generate a large set of models that are then used to compute the anisotropic U_n matrices.

Since TLS modeling is based on a linearity approximation (Urzhumtsev *et al.*, 2013) one may expect a significant
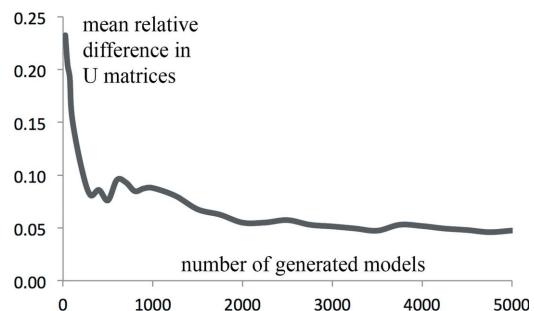

Figure 1
Comparison of the matrices U_{TLS} calculated for an atomic model analytically (see the text for details) and the matrices calculated from the ensemble models ($U_{ensemble}$) as a function of the number of models generated. The mean relative difference between the elements of the two sets of matrices is defined as $\left(\sum_n \left||U_{TLS}| + |U_{ensemble}|\right|\right)^{-1} \times \left(2\sum_n |U_{TLS} - U_{ensemble}|\right)$ where the sums are calculated over all atoms and over all six elements of each pair of U matrices.

Table 1
Matrix U calculated analytically for two atoms (see the text for details) from the TLS matrices (U_{TLS}) and numerically from an ensemble of 20 000 models ($U_{ensemble}$) for different libration amplitude dz.

Matrix elements U_{zz}, U_{xz}, U_{yz} are always equal to 0 and not shown. The last column shows the ratio of the maximal difference between respective U_{TLS} and $U_{ensemble}$ elements compared to the maximal U_{TLS} element. Two lines for $dz = 0.10$ stand for two independent runs of the random model generation.

dz (rad/°)	U_{TLS} (U_{xx}, U_{yy}, U_{xy})	$U_{ensemble}$ (U_{xx}, U_{yy}, U_{xy})	Δ_{max}/U_{max}
0.05/2.9	0.00250, 0.00063, −0.00125	0.00250, 0.00063, −0.00125	0
0.07/4.0	0.00490, 0.00123, −0.00245	0.00487, 0.00122, −0.00242	0.006
0.09/5.2	0.00810, 0.00202, −0.00405	0.00791, 0.00201, −0.00394	0.023
0.10/5.7	0.01000, 0.00250, −0.00500	0.00993, 0.00254, −0.00496	0.007
0.10/5.7	0.01000, 0.00250, −0.00500	0.00993, 0.00253, −0.00494	0.007
0.15/8.6	0.02250, 0.00562, −0.01125	0.02207, 0.00567, −0.01079	0.020
0.20/11.5	0.04000, 0.01000, −0.02000	0.03802, 0.01025, −0.01861	0.049
0.25/14.3	0.06250, 0.01562, −0.03125	0.05970, 0.01664, −0.02858	0.045
0.30/17.2	0.09000, 0.02250, −0.04500	0.08204, 0.02415, −0.03909	0.088
0.50/28.6	0.25000, 0.06250, −0.12500	0.20432, 0.07408, −0.08651	0.183
0.70/40.1	0.49000, 0.12250, −0.24500	0.32987, 0.15339, −0.11880	0.327
0.90/51.6	0.81000, 0.20250, −0.40500	0.44468, 0.25448, −0.12348	0.451

difference between the matrices U_n calculated analytically using TLS matrices and those calculated directly by (4) if the libration amplitudes are large. As mentioned in §6.1 of Urzhumtsev *et al.* (2015), large values for the vibration and libration amplitudes are not physically meaningful.

Table 1 shows matrices $U_1 = U_2$ for an artificial example of two atoms with the coordinates (0, 0, 0) and (1, 2, 3) when the only motion applied was libration around the axis parallel to **Oz** that passes through the center of mass of this system taken as the TLS origin. In the basis [M] the coordinates of these atoms are equal to (−0.5, −1.0, −1.5) and (0.5, 1.0, 1.5), respectively. The matrices were obtained using the two methods discussed above applying different amplitudes dz (for definitions see Urzhumtsev *et al.*, 2015). As discussed previously (§2.3 in Urzhumtsev *et al.*, 2013, and references therein), the discrepancy between two corresponding matrices is significant when the libration amplitude becomes larger than approximately 0.10–0.15 rad (6–9°).

To investigate the number of models sufficient to reproduce U_{TLS} by $U_{ensemble}$, we took the C_A atoms from fragment A6–A61 of protein G IgG-binding domain III model (PDB code 2igd) and fitted TLS matrices to individual anisotropic U_n of

this model using the *phenix.tls* tool. Then we calculated U_{TLS} for each atom of the model and independently generated a set of random models from which we calculated $U_{ensemble}$ to compare them with U_{TLS}. Fig. 1 shows the mean relative difference between the sets of the U elements as a function of the number of generated models.

Acknowledgements

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Grosse-Kunstleve, R. W., Sauter, N. K., Moriarty, N. W. & Adams, P. D. (2002). *J. Appl. Cryst.* **35**, 126–136.
Urzhumtsev, A., Afonine, P. V. & Adams, P. D. (2013). *Crystallogr. Rev.* **19**, 230–270.
Urzhumtsev, A., Afonine, P. V., Van Benschoten, A. H., Fraser, J. S. & Adams, P. D. (2015). *Acta Cryst.* D**71**, 1668–1683.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.

From deep TLS validation to ensembles of atomic models built from elemental motions

Alexandre Urzhumtsev,[a,b]* Pavel V. Afonine,[c] Andrew H. Van Benschoten,[d] James S. Fraser[d] and Paul D. Adams[c,e]

[a]Centre for Integrative Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS–INSERM–UdS, 1 Rue Laurent Fries, BP 10142, 67404 Illkirch, France, [b]Faculté des Sciences et Technologies, Université de Lorraine, BP 239, 54506 Vandoeuvre-les-Nancy, France, [c]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA, [d]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA 94158, USA, and [e]Department of Bioengineering, University of California Berkeley, Berkeley, CA 94720, USA. *Correspondence e-mail: sacha@igbmc.fr

The translation–libration–screw model first introduced by Cruickshank, Schomaker and Trueblood describes the concerted motions of atomic groups. Using TLS models can improve the agreement between calculated and experimental diffraction data. Because the **T**, **L** and **S** matrices describe a combination of atomic vibrations and librations, TLS models can also potentially shed light on molecular mechanisms involving correlated motions. However, this use of TLS models in mechanistic studies is hampered by the difficulties in translating the results of refinement into molecular movement or a structural ensemble. To convert the matrices into a constituent molecular movement, the matrix elements must satisfy several conditions. Refining the **T**, **L** and **S** matrix elements as independent parameters without taking these conditions into account may result in matrices that do not represent concerted molecular movements. Here, a mathematical framework and the computational tools to analyze TLS matrices, resulting in either explicit decomposition into descriptions of the underlying motions or a report of broken conditions, are described. The description of valid underlying motions can then be output as a structural ensemble. All methods are implemented as part of the *PHENIX* project.

1. Introduction

1.1. Independent and concerted molecular motions

It is currently difficult to derive a structural basis for concerted molecular motions from the models emerging from macromolecular crystallography, which describe each atom with a central position $\mathbf{r}_0$ and additional displacement parameters. Small-magnitude disorder (particularly thermal motion) can be captured by the Debye–Waller factor, which reflects the probability of an atom moving from its central position by a certain distance. If a model includes this approximation, the contribution of each atom to the structure factor (h, k, l) must be scaled by

$$\exp[-2\pi^2\mathbf{h}^\tau\mathbf{O}^{-1}\mathbf{U}_{\text{Cart}}(\mathbf{O}^{-1})^\tau\mathbf{h}], \qquad (1)$$

(see, for example, Grosse-Kunstleve & Adams, 2002 and references therein). Here, $\mathbf{O}$ is the orthogonalization matrix for the given crystal, $\mathbf{h}$ is the column vector of integer indices (h, k, l), $\mathbf{U}_{\text{Cart}}$ is an atomic displacement parameter (ADP) and the superscript τ stands for the matrix and vector transpose operation (here and in the following). [In Grosse-Kunstleve & Adams (2002) the orthogonalization matrix is defined as $\mathbf{A}$;
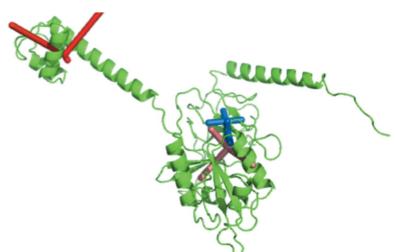
here, this letter is reserved for the matrix in the development of U_{Cart}, following Tickle & Moss (1999).] The symmetric positive definite matrix U_{Cart} is defined by the average atomic shifts (and their correlations) along each coordinate axis. The matrix U_{Cart} varies between atoms and is diagonal (with equal elements) for atoms that are assumed to be moving isotropically.

U_{Cart} can accumulate contributions from several different sources, including overall crystal anisotropy (U_{cryst}), various concerted motions (U_{group}) and independent displacement of individual atoms (U_{local}) (see, for example, Dunitz & White, 1973; Prince & Finger, 1973; Johnson, 1980; Sheriff & Hendrickson, 1987; Murshudov et al., 1999; Winn et al., 2001; , Dauter et al., 2012; Afonine et al., 2013).

Concerted motion contributing to U_{group} can be modelled by the translation–libration–screw approximation (TLS) introduced by Cruickshank (1956) and Schomaker & Trueblood (1968) and developed further in a number of publications, for example Johnson (1970), Scheringer (1973), Howlin et al. (1989, 1993), Kuriyan & Weis (1991), Schomaker & Trueblood (1998), Tickle & Moss (1999), Murshudov et al. (1999), Winn et al. (2001, 2003) and Painter & Merritt (2005, 2006a,b). This approximation is of special interest to structural biologists for two reasons. Firstly, TLS characterizes the anisotropic mobility of atomic groups and can provide insight into molecular mechanism. Secondly, it simplifies the crystallographic model by reducing the number of parameters while simultaneously providing a more realistic description of atomic displacements.

A common misconception of TLS parametrization is that its sole merit is to provide an economical method of accounting for anisotropic motions at low resolution. In fact, TLS parameterization can be useful regardless of the resolution of the available diffraction data. TLS has been successfully used to analyze functionally important molecular motions on several occasions (Kuriyan & Weis, 1991; Harris et al., 1992; Šali et al., 1992; Wilson & Brunger, 2000; Raaijmakers et al., 2001; Yousef et al., 2002; Papiz et al., 2003; Chaudhry et al., 2004), demonstrating that this approximation can provide critical structural information. However, the use of TLS models to derive functional insights is limited by the difficulty in analyzing the resulting motions. Although analysis of the resulting anisotropic displacement parameters is possible in some programs (Howlin et al., 1993; Painter & Merritt, 2005), decomposing TLS models into structural ensembles comprised of many atomic models might enable more straightforward comparisons to other data sets, particularly in the case of diffuse X-ray scattering (Van Benschoten et al., 2015). The major goal of this work is to develop an approach for translating TLS matrices into descriptions of corresponding molecular motions in terms of rotations and translations. In turn, this allows the validation of TLS parameters and the generation of structural ensembles. The latter will enable the broader use of TLS refinement for discovering and validating concerted molecular motions. In accomplishing this goal, we encountered several complications that suggest revisiting the fundamental processes of TLS refinement.

1.2. TLS model

Since the displacement of a rigid group of atoms is a composition of translation and rotation (see, for example, Goldstein, 1950), Schomaker & Trueblood (1968) presented the matrices $U_{group,n}$ for the concerted motion of a group of atoms $n = 1, 2, \ldots N$ as a sum,

$$U_{group,n} = T + A_n L A_n^\tau + A_n S + S^\tau A_n^\tau. \qquad (2)$$

The antisymmetric matrices A_n are functions of the Cartesian coordinates (x_n, y_n, z_n) of atom n

$$A_n = \begin{pmatrix} 0 & z_n & -y_n \\ -z_n & 0 & x_n \\ y_n & -x_n & 0 \end{pmatrix}. \qquad (3)$$

Matrix S and the symmetric matrices T and L are common to all atoms within each rigid group. L describes librations (oscillating rotations) around three mutually orthogonal rotation axes. T describes apparent translations of the atomic group (the term 'vibrations' might actually be more appropriate for random translations around a central position). S describes screw motions, i.e. the combination of librations and vibrations. We use the term 'apparent translation' because matrix T may have an additional contribution from librations as discussed in §2.

Thus, explicit information about atomic movement can be encoded into TLS matrices to produce inexplicit descriptors of motion. Both frameworks have merit: explicit description allows a straightforward interpretation and analysis of the motions, while the inexplicit TLS formalism provides a simpler framework for calculating structure factors. However, it is important to remember that TLS parameterization always arises from explicit atomic movement; thus, the TLS matrices should obey certain restrictions in order to be decomposed into structural ensembles representing concerted physical motions. Current refinement programs treat elements of the TLS matrices as independent variables with a constraint on the trace of the matrix S [tr(S); as discussed in §4] and post-refinement enforcement that the resulting $U_{group,n}$ be non-negative definite (Winn et al., 2001). As demonstrated below, enforcing $U_{group,n}$ to be non-negative definite is not sufficient to guarantee that the refined TLS matrices are still consistent with an underlying physical model of concerted motion.

Previously, Zucker et al. (2010) analyzed all PDB entries containing TLS descriptions and suggested tools to validate the TLS parameters. However, this analysis focused exclusively on the ADP smoothness between neighbouring TLS groups. Failure to enforce all conditions on the individual components of $U_{group,n}$, i.e. on the TLS matrices, may result in matrices that invalidate the TLS model. Using the methods and tools presented in this manuscript, we analyzed all structures from the PDB (Bernstein et al., 1977; Berman et al., 2000; about 105 000 entries, 25 000 of which contain TLS models, with a total of 200 000 sets of matrices). Our results demonstrate that significant issues are present in current TLS implementations. A third of the analyzed structures contain T or L matrices that are non-positive semidefinite and another

Table 1
Number of PDB entries in which at least one of the physical conditions on TLS matrices is broken.

The statistics are shown for the matrices in the PDB (25 904 entries with TLS matrices from a total number of 106 761 entries as of March 2015) with the default condition tr($\mathbf{S}$) = 0 (upper line) and with the optimal choice of the diagonal $\mathbf{S}$ elements whenever possible as described in §§3 and 4 (bottom line). The conditions are, from left to right: matrices $\mathbf{T}$ and $\mathbf{L}$ are positive semidefinite ($\mathbf{T} \geq 0$ and $\mathbf{L} \geq 0$); an absence of libration around one of the axes requires the corresponding elements of the $\mathbf{S}$ matrix to be equal to 0 ($s = 0$ and $w = 0$); matrix $\mathbf{T}$ is positive semidefinite after the contribution owing to the displacement of libration axes is removed ($\mathbf{T}_C \geq 0$); elements of the $\mathbf{S}$ matrix are limited by the corresponding elements of the $\mathbf{T}$ and $\mathbf{L}$ matrices according to the Cauchy conditions ($\mathbf{S} \leq \mathbf{TL}$); the residual V matrix is positive semidefinite ($\mathbf{V} \geq 0$). The column ($\mathbf{V} \geq 0$) includes all conditions from §§4.3 and 4.4. When one of the conditions was broken further conditions were not checked.

| Mode | Total No. of PDB entries | Total No. of TLS | Conditions broken | | | | | Total No. of TLS broken | Total No. of TLS OK | Total No. of PDB entries broken |
			$\mathbf{T} \geq 0$ and $\mathbf{L} \geq 0$	$s = 0$ and $w = 0$	$\mathbf{T}_C \geq 0$	$\mathbf{S} \leq \mathbf{TL}$	$\mathbf{V} \geq 0$			
$t_S = 0$	25904	203261	71362	3104	52254	n/a	10492	137212	66049	22707
Best t_S	25904	203261	71362	3104	52255	133	3776	130630	72631	22201

third (Table 1) cannot describe libration–vibration correlated motions owing to the reasons discussed in §§2–5. Some of these errors (but not all) are trivial to fix, *e.g.* correcting marginally negative eigenvalues of $\mathbf{T}$ and $\mathbf{L}$ or modifying the trace of $\mathbf{S}$ (examples are given in §6 and in Table 1).

1.3. On the physical meaning and use of TLS

Efforts to constrain TLS parameters to keep them physically meaningful have been discussed previously (Winn *et al.*, 2001; Painter & Merritt, 2006a). It is universally accepted that B values need to be positive, occupancies must range between 0 and 1 and atomic coordinates should define model geometry in accordance with chemical knowledge. Similarly, provided that the TLS groups have been selected adequately, the TLS parameters describing the anisotropic harmonic motion of atomic groups (Schomaker & Trueblood, 1968) should be physically meaningful, otherwise TLS modelling may not be considered to be applicable. One such condition, but not the only one, is that the $\mathbf{T}$ and $\mathbf{L}$ matrices are positive semidefinite.

While calculating TLS matrices from corresponding libration and vibration parameters is rather straightforward (§2), the inverse procedure is less trivial. As discussed previously (Johnson, 1970; Scheringer, 1973; Tickle & Moss, 1999), the problem itself is poorly posed since the same set of diffraction data (and consequently the same set of TLS matrices) may correspond to different motions of the contributing atoms or atomic groups. Moreover, there are computational difficulties if all the conditions on the matrices have not been considered (§§3–5).

The set of TLS matrices corresponding to physically possible combinations of motions is obviously smaller than the set of all TLS matrices. Since restricting the parameter space of any function may inadvertently exclude a number of deep minima, including the global minimum, structural refinement that imposes conditions on TLS matrices may result in higher R factors than if these conditions were ignored. Since TLS modelling is an approximation to the true molecular motions that strongly depends on the assignment of TLS groups, lower R factors as result of using TLS may not always be indicative of this model being decomposable into a valid macromolecular motion.

1.4. Summary of the presented work

In this article, we address the following points.

(i) We describe an algorithm (Fig. 1) that interprets the TLS matrices in terms of parameters of the corresponding motions. This includes the direction of the principal axes of vibration and libration, the corresponding root-mean-square displace-
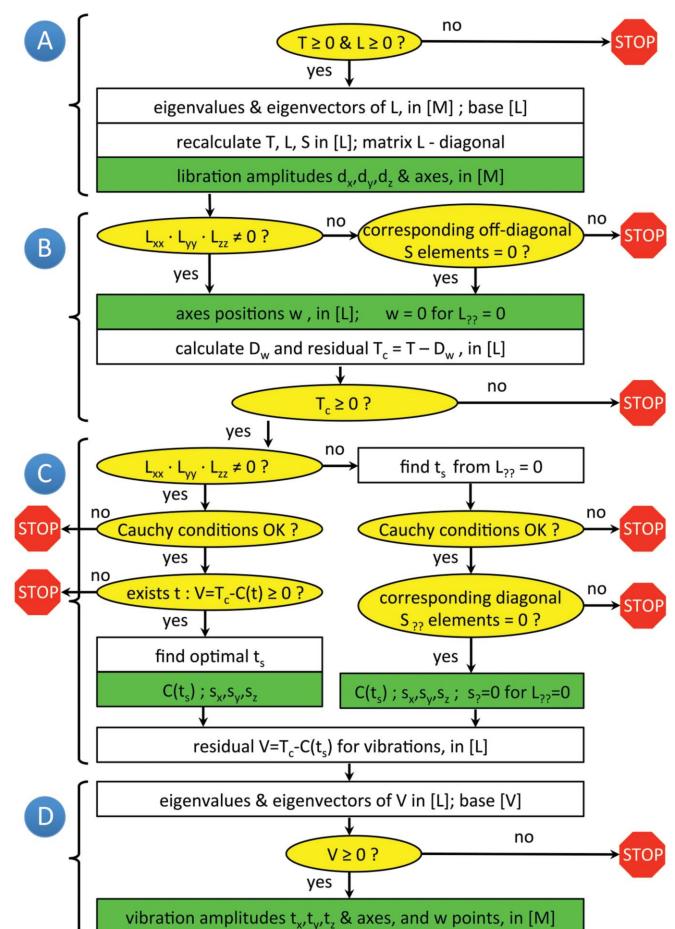

Figure 1
General flowchart of the TLS decomposition into libration and vibration composite motions. Yellow ellipses indicate conditions to be verified. Green rectangles indicate the output parameters of the composite motions. The letters *A*–*D* indicate different steps of the procedure as described in the text.

ments and the position of the libration axes, as well as the correlations between vibration and libration displacements.

(ii) We present a complete list of conditions that must be fulfilled to make the aforementioned TLS decomposition possible; this includes widely known conditions (*e.g.* **T** and **L** must be positive semidefinite) as well as a number of less trivial conditions that to the best of our knowledge have not been previously discussed.

(iii) We describe the calculation protocols in a ready-to-program style so that they can be implemented in existing or future software. Most of the calculations described in the manuscript are straightforward; less trivial expressions and proofs can be found in Appendix *A* as well as in the review by Urzhumtsev *et al.* (2013).

(iv) We implemented the described algorithms in the open-source *Computational Crystallography Toolbox* (*cctbx*; Grosse-Kunstleve *et al.*, 2002). We also made two end-user applications available in the *PHENIX* suite (Adams *et al.*, 2010): *phenix.tls_analysis* for the analysis and validation of refined TLS matrices and their underlying motions and *phenix.tls_as_xyz* for generating ensembles of structures consistent with TLS matrices.

(v) We applied these programs to all PDB entries containing TLS matrices. We discovered that the majority of these matrices cannot describe motions. In a number of cases a marginal modification of the TLS matrices can correct the errors.

(vi) We used *phenix.tls_as_xyz* to generate a predicted structural ensemble for the calculation of X-ray diffuse scattering from the glycerophosphodiesterase GpdQ (Van Benschoten *et al.*, 2015).

2. Calculating TLS matrices from elemental motions

This section provides a step-by-step protocol for calculating TLS matrices from the parameters of the composite vibrations and librations. Inverting this scheme provides a method of extracting libration/vibration parameters from the TLS matrices.

2.1. Constructing TLS matrices from the parameters of the libration and vibration

The matrices in (2) depend on the basis in which the atomic coordinates are given. We use an index in square brackets to indicate which basis is used. Let the atoms be given in some basis denoted [M]; for example, it may be the basis corresponding to the model deposited in the PDB. Even if a rigid group is involved in several simultaneous motions (assuming that the amplitudes of these motions are relatively small and the motions are harmonic), the total motion can be described by a libration around three axes $\mathbf{l}_x$, $\mathbf{l}_y$, $\mathbf{l}_z$ that are mutually orthogonal and by a vibration along three other mutually orthogonal axes, $\mathbf{v}_x$, $\mathbf{v}_y$, $\mathbf{v}_z$. These triplets of axes form the other two bases, [L] and [V].

In (2) the matrix **T** is a sum of several components. In the absence of librations (that is, matrices **L** and **S** are zero) it is equal to the contribution **V** arising from pure vibrations. In the basis [V] this matrix is diagonal,

$$\mathbf{V}_{[V]} = \begin{pmatrix} \langle t_x^2 \rangle & 0 & 0 \\ 0 & \langle t_y^2 \rangle & 0 \\ 0 & 0 & \langle t_z^2 \rangle \end{pmatrix}. \tag{4}$$

Here, $\langle t_x^2 \rangle$, $\langle t_y^2 \rangle$, $\langle t_z^2 \rangle$ are the corresponding squared root-mean-square deviations (r.m.s.d.s) along the principal vibration axes $\mathbf{v}_x$, $\mathbf{v}_y$, $\mathbf{v}_z$ and are expressed in Å^2. If there are librations, the matrix **L** is always diagonal in the basis [L],

$$\mathbf{L}_{[L]} = \begin{pmatrix} \langle d_x^2 \rangle & 0 & 0 \\ 0 & \langle d_y^2 \rangle & 0 \\ 0 & 0 & \langle d_z^2 \rangle \end{pmatrix}. \tag{5}$$

Here, $\langle d_x^2 \rangle$, $\langle d_y^2 \rangle$, $\langle d_z^2 \rangle$ are the squared r.m.s.d.s of the vibration angles expressed in squared radians; for small deviations they are numerically equal to the squared r.m.s.d.s of points at a unit distance from the corresponding axes.

In reality, the principal vibration and libration axes are not parallel to each other; practically, it is convenient to express the matrices in a common basis. Basis [L] is more convenient for this since in this basis the elements of **S** (see below) are easily expressed through geometric parameters of librations. Matrix **V** in this basis is no longer diagonal but is instead equal to

$$\mathbf{V}_{[L]} = \mathbf{R}_{VL}^\tau \mathbf{V}_{[L]} \mathbf{R}_{VL}. \tag{6}$$

Here, $\mathbf{R}_{VL}$ is the transition matrix that describes the rotation superposing the vectors $\mathbf{v}_x$, $\mathbf{v}_y$, $\mathbf{v}_z$ with the vectors $\mathbf{l}_x$, $\mathbf{l}_y$, $\mathbf{l}_z$ (Appendix *A*). Frequently, vibration and libration motions are not independent but instead are correlated to form screw rotations. It is convenient to characterize screw rotations by the parameters s_x, s_y, s_z: for a screw rotation by d_z radians around an axis parallel to $\mathbf{l}_z$ each atom is shifted by s_z Å along this axis. A similar definition is used for the other two parameters. If the axes pass through the origin, such a correlation generates an additional contribution $\mathbf{C}_{[L]}$ to the **T** matrix that arises from screw motions,

$$\mathbf{T}_{C[L]} = \mathbf{V}_{[L]} + \mathbf{C}_{[L]} = \mathbf{V}_{[L]} + \begin{pmatrix} s_x^2 \langle d_x^2 \rangle & 0 & 0 \\ 0 & s_y^2 \langle d_y^2 \rangle & 0 \\ 0 & 0 & s_z^2 \langle d_z^2 \rangle \end{pmatrix}, \tag{7}$$

and also results in a nonzero **S** matrix,

$$\mathbf{S}_{[L]} = \begin{pmatrix} s_x \langle d_x^2 \rangle & 0 & 0 \\ 0 & s_y \langle d_y^2 \rangle & 0 \\ 0 & 0 & s_z \langle d_z^2 \rangle \end{pmatrix}. \tag{8}$$

Finally, the principal libration axes do not necessarily pass through the origin, or even have a common point (*i.e.* they may not intersect). If they pass through the points $\mathbf{w}_{[L]}^{lx} = (w_x^{lx}, w_y^{lx}, w_z^{lx})$, $\mathbf{w}_{[L]}^{ly} = (w_x^{ly}, w_y^{ly}, w_z^{ly})$, $\mathbf{w}_{[L]}^{lz} = (w_x^{lz}, w_y^{lz}, w_z^{lz})$, respectively, this generates an additional component to the **T** matrix,

$$\mathbf{T}_{[L]} = \mathbf{T}_{C[L]} + \mathbf{D}_{W[L]}, \tag{9}$$

where

research papers

$$\mathbf{D}_{\mathrm{W[L]}} =$$

$$\begin{bmatrix} (w_z^i)^2 \langle d_y^2 \rangle + (w_y^k)^2 \langle d_z^2 \rangle & -w_x^k w_y^k \langle d_z^2 \rangle & -w_x^i w_z^i \langle d_y^2 \rangle \\ -w_x^k w_y^k \langle d_z^2 \rangle & (w_z^i)^2 \langle d_x^2 \rangle + (w_x^k)^2 \langle d_z^2 \rangle & -w_y^i w_z^i \langle d_x^2 \rangle \\ -w_x^i w_z^i \langle d_y^2 \rangle & -w_y^i w_z^i \langle d_x^2 \rangle & (w_y^i)^2 \langle d_x^2 \rangle + (w_x^i)^2 \langle d_y^2 \rangle \end{bmatrix}.$$

$$(10)$$

Taking into account both the screw motion and the position of the libration axes, the matrix $\mathbf{S}$ becomes

$$\mathbf{S}_{[\mathrm{L}]} = \begin{pmatrix} s_x \langle d_x^2 \rangle & w_z^i \langle d_x^2 \rangle & -w_y^i \langle d_x^2 \rangle \\ -w_z^j \langle d_y^2 \rangle & s_y \langle d_y^2 \rangle & w_x^j \langle d_y^2 \rangle \\ w_y^k \langle d_z^2 \rangle & -w_x^k \langle d_z^2 \rangle & s_z \langle d_z^2 \rangle \end{pmatrix}. \quad (11)$$

Finally, the matrices in the original basis [M] where they are reported together with the atomic coordinates are obtained from $\mathbf{L}_{[\mathrm{L}]}$ (5), $\mathbf{T}_{[\mathrm{L}]}$ (9), $\mathbf{S}_{[\mathrm{L}]}$ (11) as

$$\mathbf{L}_{[\mathrm{M}]} = \mathbf{R}_{\mathrm{ML}} \mathbf{L}_{[\mathrm{L}]} \mathbf{R}_{\mathrm{ML}}^{\tau},$$
$$\mathbf{T}_{[\mathrm{M}]} = \mathbf{R}_{\mathrm{ML}} \mathbf{T}_{[\mathrm{L}]} \mathbf{R}_{\mathrm{ML}}^{\tau},$$
$$\mathbf{S}_{[\mathrm{M}]} = \mathbf{R}_{\mathrm{ML}} \mathbf{S}_{[\mathrm{L}]} \mathbf{R}_{\mathrm{ML}}^{\tau}. \quad (12)$$

Here, $\mathbf{R}_{\mathrm{ML}}$ is the transition matrix from the basis [M] to the basis [L] (Appendix A).

2.2. Molecular basis and centre of reaction

The TLS matrices also depend on the choice of the origin. Clearly, the coordinates of the position of the libration axes change as function of the origin. Usually, the origin is taken to be the centre of mass of the atomic group or the point where the mean atomic displacements are similar in magnitude to each other owing to librations around each of the principal axes. This second point is called the centre of diffusion (Brenner, 1967) or the centre of reaction (Tickle & Moss, 1999). Choosing the origin at the centre of reaction minimizes the trace of $\mathbf{T}$ and makes $\mathbf{S}$ symmetric (Brenner, 1967; Tickle & Moss, 1999; Urzhumtsev *et al.*, 2013). Shifting from one origin to another changes $\mathbf{T}$ and $\mathbf{S}$ but does not change $\mathbf{L}$ and does not modify the algorithm of the search for the composite motions. In the following, we consider the matrices to be in their original basis (for example, as they are defined in the PDB).

3. Calculating elemental motions from TLS matrices: libration axes

This section provides a step-by-step explanation of the inverse problem, *i.e.* calculating the vibration and libration axes and the corresponding r.m.s.d.s, the position of the libration axes and the parameters describing the correlations between librations and vibrations from given TLS matrices.

3.1. Diagonalization of the L matrix ([L] basis; step *A*)

Suppose that we know the elements of the matrices (12) in the basis [M]. By construction, the matrices $\mathbf{T}$ and $\mathbf{L}$ should be positive semidefinite (Appendix B) and symmetric, $T_{[\mathrm{M}]xy} = T_{[\mathrm{M}]yx}$, $T_{[\mathrm{M}]xz} = T_{[\mathrm{M}]zx}$, $T_{[\mathrm{M}]yz} = T_{[\mathrm{M}]zy}$ and $L_{[\mathrm{M}]xy} = L_{[\mathrm{M}]yx}$, $L_{[\mathrm{M}]xz} = L_{[\mathrm{M}]zx}$, $L_{[\mathrm{M}]yz} = L_{[\mathrm{M}]zy}$. These properties hold for any

rotation of the coordinate system, *i.e.* in any Cartesian basis; this is important for further analysis of the $\mathbf{T}$ matrices.

We start the procedure from the matrix $\mathbf{L}_{[\mathrm{M}]}$, which depends only on the libration parameters. The principal libration axes correspond to its three mutually orthogonal eigenvectors. Firstly, we search for the corresponding eigenvalues $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3$, which must be non-negative (see equation 5; eigenvalues do not change with the coordinate system). Let $\mathbf{l}_1, \mathbf{l}_2, \mathbf{l}_3$ be the corresponding normalized eigenvectors from which we construct a new basis [L] as

$$\mathbf{l}_x = \pm\mathbf{l}_1, \quad \mathbf{l}_y = \mathbf{l}_2, \quad \mathbf{l}_z = \mathbf{l}_3. \quad (13)$$

The appropriate sign for $\mathbf{l}_x$ is chosen so that the vectors in (13) form a right-hand triad; for example, one can take $\mathbf{l}_x = \mathbf{l}_y \times \mathbf{l}_z$ which guarantees such a condition. The TLS matrices in the [L] basis are

$$\mathbf{L}_{[\mathrm{L}]} = \mathbf{R}_{\mathrm{ML}}^{\tau} \mathbf{L}_{[\mathrm{M}]} \mathbf{R}_{\mathrm{ML}},$$
$$\mathbf{T}_{[\mathrm{L}]} = \mathbf{R}_{\mathrm{ML}}^{\tau} \mathbf{T}_{[\mathrm{M}]} \mathbf{R}_{\mathrm{ML}},$$
$$\mathbf{S}_{[\mathrm{L}]} = \mathbf{R}_{\mathrm{ML}}^{\tau} \mathbf{S}_{[\mathrm{M}]} \mathbf{R}_{\mathrm{ML}}, \quad (14)$$

where $\mathbf{R}_{\mathrm{ML}}$ is the transition matrix from basis [M] into basis [L] (Appendix A). In this new basis, matrix $\mathbf{L}_{[\mathrm{L}]}$ is diagonal with elements $L_{[\mathrm{L}]xx} = \lambda_1$, $L_{[\mathrm{L}]yy} = \lambda_2$, $L_{[\mathrm{L}]zz} = \lambda_3$, giving the estimates $\langle d_x^2 \rangle = L_{[\mathrm{L}]xx}$, $\langle d_y^2 \rangle = L_{[\mathrm{L}]yy}$, $\langle d_z^2 \rangle = L_{[\mathrm{L}]zz}$ of the squared libration amplitudes around the three principal libration axes.

3.2. Position of the libration axes in the [L] basis (step *B*)

In the basis [L] the libration axes are parallel to the co-ordinate axes but do not necessarily coincide with them. Let them pass through some points $\mathbf{w}^{lx}$, $\mathbf{w}^{ly}$, $\mathbf{w}^{lz}$, respectively, that must be identified. Using (11), we calculate the coordinates of these points as

$$\mathbf{w}_{[\mathrm{L}]y}^{lx} = -\frac{S_{[\mathrm{L}]xz}}{L_{[\mathrm{L}]xx}}, \quad \mathbf{w}_{[\mathrm{L}]z}^{lx} = \frac{S_{[\mathrm{L}]xy}}{L_{[\mathrm{L}]xx}},$$

$$\mathbf{w}_{[\mathrm{L}]x}^{ly} = \frac{S_{[\mathrm{L}]yz}}{L_{[\mathrm{L}]yy}}, \quad \mathbf{w}_{[\mathrm{L}]z}^{ly} = -\frac{S_{[\mathrm{L}]yx}}{L_{[\mathrm{L}]yy}},$$

$$\mathbf{w}_{[\mathrm{L}]x}^{lz} = -\frac{S_{[\mathrm{L}]zy}}{L_{[\mathrm{L}]zz}}, \quad \mathbf{w}_{[\mathrm{L}]y}^{lz} = \frac{S_{[\mathrm{L}]zx}}{L_{[\mathrm{L}]zz}}. \quad (15)$$

A zero value for any denominator in (15) means that there is no rotation around the corresponding axis; in this case, the two corresponding numerator values must also be equal to zero and thus assign zero values to the corresponding coordinates in (15). Otherwise, the input matrices are incompatible and the procedure must stop (Appendix B). The x component of $\mathbf{w}^{lx}$, the y component of $\mathbf{w}^{ly}$ and the z component of $\mathbf{w}^{lz}$ in the basis [L] can be any values. For presentation purposes, it might be useful to assign them as

$$\mathbf{w}_{[\mathrm{L}]x}^{lx} = \tfrac{1}{2}(\mathbf{w}_{[\mathrm{L}]x}^{ly} + \mathbf{w}_{[\mathrm{L}]x}^{lz}),$$
$$\mathbf{w}_{[\mathrm{L}]y}^{ly} = \tfrac{1}{2}(\mathbf{w}_{[\mathrm{L}]y}^{lx} + \mathbf{w}_{[\mathrm{L}]y}^{lz}),$$
$$\mathbf{w}_{[\mathrm{L}]z}^{lz} = \tfrac{1}{2}(\mathbf{w}_{[\mathrm{L}]z}^{lx} + \mathbf{w}_{[\mathrm{L}]z}^{ly}) \quad (16)$$

that will position each of these points in the middle of the two other axes. This choice also reduces eventual rounding errors.

Knowing the positions (15 and 16) of the libration axes and elements of $\mathbf{L}_{[L]}$, we can calculate the contribution $\mathbf{D}_{W[L]}$ (10) from an apparent translation owing to the displacement of the libration axes from the origin. Then, by inverting (9) we can calculate the residual matrix $\mathbf{T}_{C[L]}$ after removal of this contribution,

$$\mathbf{T}_{C[L]} = \mathbf{T}_{[L]} - \mathbf{D}_{W[L]}. \tag{17}$$

Matrix (17) must be positive semidefinite (Appendix B) as it is a sum (7) of two positive semidefinite matrices. Matrices $\mathbf{S}_{[L]}$ and $\mathbf{L}_{[L]}$ are not modified at this step.

4. Calculating elemental motions from TLS matrices: screw components (step C)

4.1. Correlation between libration and vibration and a choice of the diagonal elements of S

Next, we use the matrices $\mathbf{L}_{[L]}$ and $\mathbf{S}_{[L]}$ to determine the screw parameters s_x, s_y, s_z, remove the screw contribution from the $\mathbf{T}_{C[L]}$ matrix using (7) and (17) and finally extract the matrix $\mathbf{V}_{[L]}$ for uncorrelated vibrations. However, there is an ambiguity in the definition of $\mathbf{S}_{[L]}$ which is apparent from the observation that the matrices $\mathbf{U}_{concerted,n}$ of individual atoms will not change if the same number t is added or removed simultaneously from all three diagonal elements of $\mathbf{S}_{[L]}$. This is usually known as indetermination of the trace of this matrix (Schomaker & Trueblood, 1968). A current practice (an illustration is provided in §6.1) is to choose t such that it minimizes the trace (rather its absolute value) of the resulting matrix,

$$\mathbf{S}_C(t) = \mathbf{S}_{[L]} - t\mathbf{I} \tag{18}$$

(where I is a unit matrix), *i.e.* minimizing vibration–libration correlation (Urzhumtsev *et al.*, 2013), or simply makes the trace equal to zero (http://www.ccp4.ac.uk/html/restrain.html; Coppens, 2006). The unconditioned minimization

$$\min_t |\mathbf{S}_C(t)| = \min_t |(S_{[L]xx} + S_{[L]yy} + S_{[L]zz}) - 3t| = 0 \tag{19}$$

gives

$$t_0 = \tfrac{1}{3}(S_{[L]xx} + S_{[L]yy} + S_{[L]zz}) = \tfrac{1}{3}\mathrm{tr}(\mathbf{S}_{[L]}). \tag{20}$$

However, this value may lead to matrices for which libration–vibration decomposition is impossible and, in particular, prohibits the generation of structural ensembles. For example, if the elements of matrix $\mathbf{S}$ and the corresponding values s_x, s_y, s_z are too large, the matrix $\mathbf{V}$ in (7) may be not positive definite for a given $\mathbf{T}_{C[L]}$. The next sections describe a procedure that defines the constraints on the diagonal elements of matrix $\mathbf{S}$ when using (18).

4.2. Cauchy–Schwarz conditions

After removing $\mathbf{D}_{W[L]}$ (17), the set of matrices $\mathbf{T}_{C[L]}$, $\mathbf{L}_{[L]}$ and the matrix $\mathbf{S}_{[L]}$ with the removed off-diagonal elements (reducing the matrix in equation 11 to the form in equation 8)

correspond to a combination of vibrations with screw rotations around the axes crossing the origin. The diagonal elements of these matrices must satisfy the Cauchy–Schwarz inequality (Appendix A),

$$
\begin{aligned}
S_{C,xx}^2 &\le T_{C[L]xx}L_{[L]xx}, \\
S_{C,yy}^2 &\le T_{C[L]yy}L_{[L]yy}, \\
S_{C,zz}^2 &\le T_{C[L]zz}L_{[L]zz},
\end{aligned} \tag{21}
$$

that in turn defines the conditions (Appendices A and B)

$$
\begin{aligned}
(S_{[L]xx} - t)^2 &\le T_{C[L]xx}L_{[L]xx}, \\
(S_{[L]yy} - t)^2 &\le T_{C[L]yy}L_{[L]yy}, \\
(S_{[L]zz} - t)^2 &\le T_{C[L]zz}L_{[L]zz}
\end{aligned} \tag{22}
$$

or

$$t_{\min,C} \le t \le t_{\max,C} \tag{23}$$

with

$$
\begin{aligned}
t_{\min,C} &= \max\{S_{[L]xx} - r_x;\, S_{[L]yy} - r_y;\, S_{[L]zz} - r_z\}, \\
t_{\max,C} &= \min\{S_{[L]xx} + r_x;\, S_{[L]yy} + r_y;\, S_{[L]zz} + r_z\}, \\
r_x &= (\tfrac{1}{2}T_{C[L]xx}L_{[L]xx})^{1/2}, \\
r_y &= (\tfrac{1}{2}T_{C[L]yy}L_{[L]yy})^{1/2}, \\
r_z &= (\tfrac{1}{2}T_{C[L]zz}L_{[L]zz})^{1/2}.
\end{aligned} \tag{24}
$$

In particular, this unambiguously defines the t value if one of the diagonal elements of the matrix $\mathbf{L}_{[L]}$ is zero so that the trace of $\mathbf{S}_{[L]}$ cannot be changed or assigned arbitrarily (see §4.4).

4.3. Positive semidefinition of the V matrix

The last condition to check is that the matrix $\mathbf{V}$ is positive semidefinite. Let us suppose that all diagonal elements of the matrix $\mathbf{L}_{[L]}$ are different from zero; §4.4 considers the alternative case. From (5), (7), (8) and (18) we find the expression for the screw contribution

$$
\begin{aligned}
\mathbf{C}_{[L]}(t) &= \begin{pmatrix} S_{C,xx}^2 L_{[L]xx}^{-1} & 0 & 0 \\ 0 & S_{C,yy}^2 L_{[L]yy}^{-1} & 0 \\ 0 & 0 & S_{C,zz}^2 L_{[L]zz}^{-1} \end{pmatrix} \\
&= \begin{bmatrix} (S_{[L]xx} - t)^2 L_{[L]xx}^{-1} & 0 & 0 \\ 0 & (S_{[L]yy} - t)^2 L_{[L]yy}^{-1} & 0 \\ 0 & 0 & (S_{[L]zz} - t)^2 L_{[L]zz}^{-1} \end{bmatrix}
\end{aligned} \tag{25}
$$

to be subtracted from matrix (17) as

$$\mathbf{V}_{[L]} = \mathbf{T}_{C[L]} - \mathbf{C}_{[L]}(t). \tag{26}$$

Matrix $\mathbf{V}_{[L]}$ is positive semidefinite along with

$$\mathbf{V}_{\Lambda} = \begin{pmatrix} v_{XX} & v_{XY} & v_{XZ} \\ v_{YX} & v_{YY} & v_{YZ} \\ v_{ZX} & v_{ZY} & v_{ZZ} \end{pmatrix} = \mathbf{\Lambda}^{\tau} \mathbf{V}_{[\mathrm{L}]} \mathbf{\Lambda}$$

$$= \mathbf{\Lambda}^{\tau} \mathbf{T}_{\mathrm{C}[\mathrm{L}]} \mathbf{\Lambda} - \mathbf{\Lambda}^{\tau} \mathbf{C}_{[\mathrm{L}]}(t) \mathbf{\Lambda} = \mathbf{T}_{\Lambda} - \mathbf{C}_{\Lambda}(t), \qquad (27)$$

where

$$\mathbf{\Lambda} = \mathbf{\Lambda}^{\tau} = \begin{pmatrix} L_{[\mathrm{L}]xx}^{1/2} & 0 & 0 \\ 0 & L_{[\mathrm{L}]yy}^{1/2} & 0 \\ 0 & 0 & L_{[\mathrm{L}]zz}^{1/2} \end{pmatrix}, \qquad (28)$$

$$\mathbf{T}_{\Lambda} = \begin{pmatrix} T_{\mathrm{C}[\mathrm{L}]xx} L_{[\mathrm{L}]xx} & T_{\mathrm{C}[\mathrm{L}]xy} L_{[\mathrm{L}]xx}^{1/2} L_{[\mathrm{L}]yy}^{1/2} & T_{\mathrm{C}[\mathrm{L}]xz} L_{[\mathrm{L}]xx}^{1/2} L_{[\mathrm{L}]zz}^{1/2} \\ T_{\mathrm{C}[\mathrm{L}]yx} L_{[\mathrm{L}]xx}^{1/2} L_{[\mathrm{L}]yy}^{1/2} & T_{\mathrm{C}[\mathrm{L}]yy} L_{[\mathrm{L}]yy} & T_{\mathrm{C}[\mathrm{L}]yz} L_{[\mathrm{L}]yy}^{1/2} L_{[\mathrm{L}]zz}^{1/2} \\ T_{\mathrm{C}[\mathrm{L}]zx} L_{[\mathrm{L}]xx}^{1/2} L_{[\mathrm{L}]zz}^{1/2} & T_{\mathrm{C}[\mathrm{L}]zy} L_{[\mathrm{L}]yy}^{1/2} L_{[\mathrm{L}]zz}^{1/2} & T_{\mathrm{C}[\mathrm{L}]zz} L_{[\mathrm{L}]zz} \end{pmatrix},$$

$$(29)$$

$$\mathbf{C}_{\Lambda}(t) = \begin{bmatrix} (S_{[\mathrm{L}]xx} - t)^2 & 0 & 0 \\ 0 & (S_{[\mathrm{L}]yy} - t)^2 & 0 \\ 0 & 0 & (S_{[\mathrm{L}]zz} - t)^2 \end{bmatrix}. \qquad (30)$$

Necessarily, all diagonal terms of (30) cannot be larger than the maximal eigenvalue τ_{max} of matrix (29), giving a necessary condition (Appendix B)

$$t_{\mathrm{min},\tau} \leq t \leq t_{\mathrm{max},\tau}$$
$$t_{\mathrm{min},\tau} = \max\{S_{[\mathrm{L}]xx}; S_{[\mathrm{L}]yy}; S_{[\mathrm{L}]zz}\} - \tau_{\mathrm{max}}^{1/2}$$
$$t_{\mathrm{max},\tau} = \min\{S_{[\mathrm{L}]xx}; S_{[\mathrm{L}]yy}; S_{[\mathrm{L}]zz}\} + \tau_{\mathrm{max}}^{1/2}. \qquad (31)$$

The condition that all diagonal terms of (30) are not larger than the minimum eigenvalue τ_{min} of (29) is sufficient but not necessary.

Matrix $\mathbf{V}_{\Lambda}$ is positive semidefinite if and only if all three of its real eigenvalues are non-negative (some of them may coincide with each other). They are the roots of the cubic characteristic equation

$$v^3 + a_{\mathrm{S}} v^2 + b_{\mathrm{S}} v + c_{\mathrm{S}} = 0, \qquad (32)$$

with the coefficients

$$a_{\mathrm{S}}(t) = -\mathrm{tr}(\mathbf{V}_{\Lambda}), \qquad (33)$$

$$b_{\mathrm{S}}(t) = \det \begin{pmatrix} v_{XX} & v_{XY} \\ v_{YX} & v_{YY} \end{pmatrix} + \det \begin{pmatrix} v_{YY} & v_{YZ} \\ v_{ZY} & v_{ZZ} \end{pmatrix}$$

$$+ \det \begin{pmatrix} v_{ZZ} & v_{ZX} \\ v_{XZ} & v_{XX} \end{pmatrix}, \qquad (34)$$

$$c_{\mathrm{S}}(t) = -\det(\mathbf{V}_{\Lambda}). \qquad (35)$$

The roots of (32) are positive if and only if the three inequalities below hold simultaneously,

$$a_{\mathrm{S}}(t) \leq 0, \quad b_{\mathrm{S}}(t) \geq 0, \quad c_{\mathrm{S}}(t) \leq 0, \qquad (36)$$

where the left parts are polynomials of order two, four and six of the parameter t, all with the unit highest-order coefficient (Appendix A). The first condition in (36) defines the interval for t values (Appendix B),

$$t_{\mathrm{min},a} = t_0 - t_a \leq t \leq t_{\mathrm{max,a}} = t_0 + t_a \qquad (37)$$

with

$$t_a = [t_0^2 + \tfrac{1}{3}\mathrm{tr}(\mathbf{T}_{\Lambda}) - \tfrac{1}{3}(S_{[\mathrm{L}]xx}^2 + S_{[\mathrm{L}]yy}^2 + S_{[\mathrm{L}]zz}^2)]^{1/2}. \qquad (38)$$

We failed to find analytical expressions corresponding to the two other inequalities. As a result, the following numerical procedure is suggested to find the best t value that is physically acceptable.

(i) Calculate the t_0 value (20).

(ii) Calculate the interval (t_{min}, t_{max}) for allowed t values as the intersection of intervals (23), (31) and (37), $t_{\mathrm{min}} = \max\{t_{\mathrm{min,C}}, t_{\mathrm{min},\tau}, t_{\mathrm{min,a}}\}$, $t_{\mathrm{max}} = \min\{t_{\mathrm{max,C}}, t_{\mathrm{max},\tau}, t_{\mathrm{max,a}}\}$; if $t_{\mathrm{min}} > t_{\mathrm{max}}$ the problem has no solution and the procedure stops (Appendix B).

(iii) If $t_{\mathrm{min}} = t_{\mathrm{max}}$ we check the conditions $b_{\mathrm{S}}(t_{\mathrm{min}}) \geq 0$, $c_{\mathrm{S}}(t_{\mathrm{min}}) \leq 0$, or the condition that $\mathbf{V}_{\Lambda}$ is positive semidefinite; if the conditions are satisfied we assign $t_{\mathrm{S}} = t_{\mathrm{min}}$, otherwise the problem has no solution and the procedure stops (Appendix B).

(iv) If $t_{\mathrm{min}} < t_{\mathrm{max}}$ we search numerically, in a fine grid, for the point t_{S} in the interval (t_{min}, t_{max}) and closest to t_0 such that $b_{\mathrm{S}}(t_{\mathrm{S}}) \geq 0$, $c_{\mathrm{S}}(t_{\mathrm{S}}) \leq 0$; if for any point of this interval at least one
of these inequalities is wrong then the procedure stops (Appendix B).

(v) We accept the value obtained at the step (iii) or (iv) as the final t_{S}.

4.4. Singular sets of rotation

When one of the $L_{[\mathrm{L}]xx}$, $L_{[\mathrm{L}]yy}$, $L_{[\mathrm{L}]zz}$ values is zero (that is, there is no rotation around the corresponding axis), straightforward use of the standard procedure including (25) becomes impossible. However, in this case the t_{S} value must be equal to $S_{[\mathrm{L}]xx}$, $S_{[\mathrm{L}]yy}$ or $S_{[\mathrm{L}]zz}$, corresponding to the axes with no rotation, making the corresponding diagonal element in (25) equal to zero and turning the corresponding inequality in (24) into an equality. For example, if $L_{[\mathrm{L}]xx} = 0$ then $t_{\mathrm{S}} = S_{[\mathrm{L}],xx}$, resulting in $C_{[\mathrm{L}]xx} = 0$. We simply need to check two other conditions in (21) and confirm that the residual matrix is positive semidefinite (for example, by calculating equation 36). If t_{S} does not satisfy these conditions, the problem has no solution (Appendix B).

4.5. Screw parameters

For the $t = t_{\mathrm{S}}$ determined above we calculate the matrix $\mathbf{S}_{\mathrm{C}}(t_{\mathrm{S}})$ (18). From this matrix we obtain the screw parameters $s_x = S_{\mathrm{C},xx} L_{[\mathrm{L}]xx}^{-1}$, $s_y = S_{\mathrm{C},yy} L_{[\mathrm{L}]yy}^{-1}$, $s_z = S_{\mathrm{C},zz} L_{[\mathrm{L}]zz}^{-1}$ for the rotation axes currently aligned with the coordinate axes of the basis [L]. If one of the $L_{[\mathrm{L}]xx}$, $L_{[\mathrm{L}]yy}$, $L_{[\mathrm{L}]zz}$ values is equal to zero, the corresponding diagonal element of $\mathbf{S}_{\mathrm{C}}$ must also be equal to zero, and we assign the corresponding screw parameter, s_x, s_y or s_z, to be zero. Otherwise, the matrices are inconsistent with each other and the procedure stops (Appendix B).

5. Calculating elemental motions from TLS matrices: vibration components (step D)

5.1. Matrix V and vibration parameters in [L] basis

For the known t_S, the matrices $\mathbf{C}_{[L]}(t_S)$ and then $\mathbf{V}_{[L]}$ are calculated using (25) and (26). The parameter values of the independent vibrations are calculated from the $\mathbf{V}_{[L]}$ matrix similarly to those for the independent librations, as we obtain them from $\mathbf{L}_{[M]}$. Firstly, we calculate the three eigenvalues $0 \leq \mu_1 \leq \mu_2 \leq \mu_3$ of matrix $\mathbf{V}_{[L]}$ (Appendix B; in practice, all of them are strictly positive). We then identify three corresponding unit eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ that are orthogonal to each other and assign

$$\mathbf{v}_x = \pm\mathbf{v}_1, \mathbf{v}_y = \mathbf{v}_2, \mathbf{v}_z = \mathbf{v}_3 \qquad (39)$$

[the sign for $\mathbf{v}_x$ is taken so that the vectors (39) form a right-hand triad]. We remind the reader that these axes define the basis [V] in which matrix $\mathbf{V}_{[V]}$ (6) is diagonal with elements $V_{[V]xx} = \mu_1$, $V_{[V]yy} = \mu_2$, $V_{[V]zz} = \mu_3$. This defines the last missing parameters, namely the values of the squared r.m.s.d.s along these axes: $\langle t_x^2 \rangle = V_{[V]xx}$, $\langle t_y^2 \rangle = V_{[V]yy}$, $\langle t_z^2 \rangle = V_{[V]zz}$.
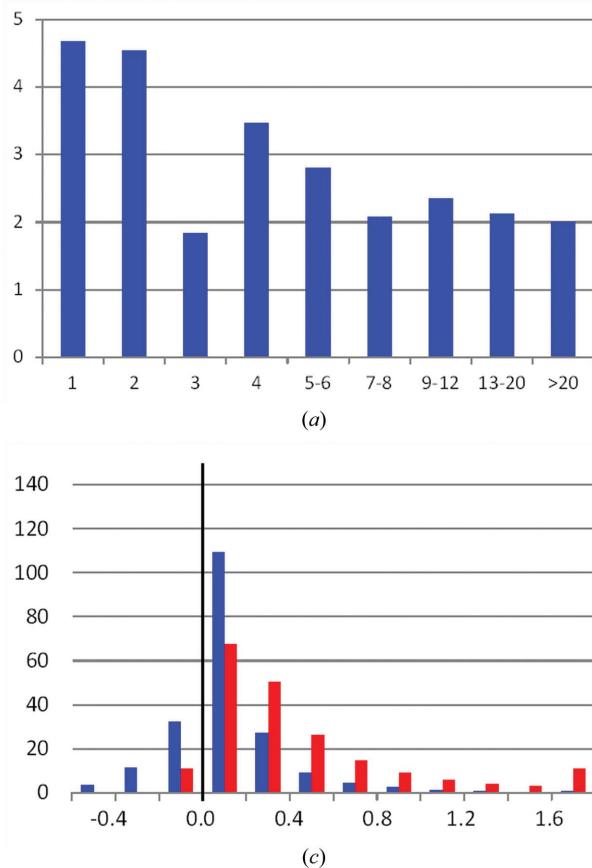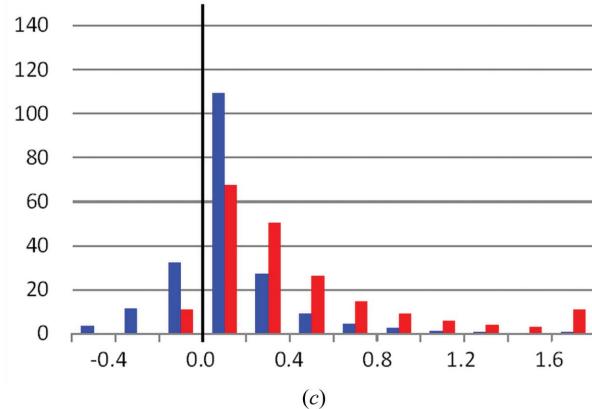
5.2. Vibration and libration axes in [M] basis

The libration and vibration amplitudes and screw parameters are independent of the choice of the basis, and the direction of the libration axes is known in the principal [M] basis. However, the directions of the uncorrelated translations $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ that were calculated in §4 and the points $\mathbf{w}_{[L]}^{lx}, \mathbf{w}_{[L]}^{ly}, \mathbf{w}_{[L]}^{lz}$ belonging to the libration axes (§3.2) are now known in the [L] basis.

To obtain the coordinates $(w_{[M]x}^{lx}, w_{[M]y}^{lx}, w_{[M]z}^{lx})$, $(w_{[M]x}^{ly}, w_{[M]y}^{ly}, w_{[M]z}^{ly})$, $(w_{[M]x}^{lz}, w_{[M]y}^{lz}, w_{[M]z}^{lz})$ of these points in the [M] basis, we apply the transformation

$$\mathbf{w}_{[M]}^{lx} = \mathbf{R}_{ML}\mathbf{w}_{[L]}^{lx}, \qquad \mathbf{w}_{[M]}^{ly} = \mathbf{R}_{ML}\mathbf{w}_{[L]}^{ly}, \qquad \mathbf{w}_{[M]}^{lz} = \mathbf{R}_{ML}\mathbf{w}_{[L]}^{lz}. \qquad (40)$$

Similarly, the vectors defining the direction of the axes $\mathbf{v}_x, \mathbf{v}_y, \mathbf{v}_z$ in the basis [M] can be obtained as

$$\mathbf{v}_{[M]x} = \mathbf{R}_{ML}\mathbf{v}_{[L]x}, \qquad \mathbf{v}_{[M]y} = \mathbf{R}_{ML}\mathbf{v}_{[L]y}, \qquad \mathbf{v}_{[M]z} = \mathbf{R}_{ML}\mathbf{v}_{[L]z}. \qquad (41)$$

This step finalizes the extraction of the parameters of the motions that correspond to the given set of TLS matrices. §6
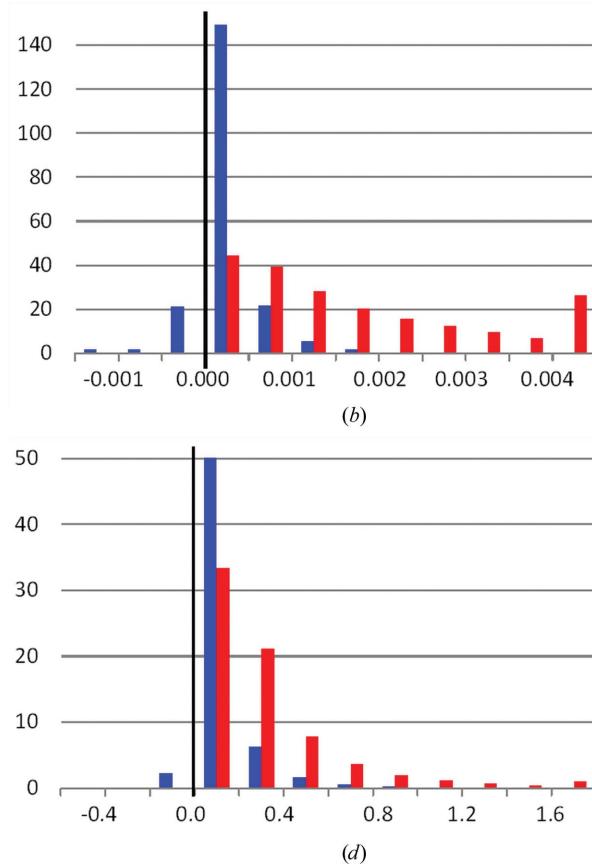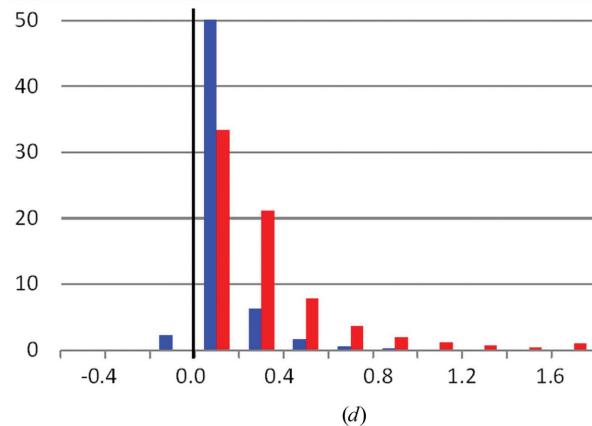

Figure 2
The number of PDB entries (in thousands) as a function of various parameters. The blue histogram in (b), (c) and (d) is for the minimum eigenvalue and the red histogram is for the maximum eigenvalue. The leftmost and rightmost bins include all cases with values less than or greater than the limits given on the axis, respectively. The eigenvalues are given in rad^2 for $\mathbf{L}$ and in Å^2 for $\mathbf{T}$. The total number of TLS groups is 203 261 for (a), (b) and (c) and about 70 000 for (d) when the matrix $\mathbf{V}$ could be calculated. (a) The number of TLS groups per entry; the largest is 283. (b) Distribution of eigenvalues of the matrix $\mathbf{L}$; the minimum eigenvalue varies from -0.285 to 0.164 and the maximum eigenvalue varies from -0.001 to 0.409. (c) Distribution of eigenvalues of the matrix $\mathbf{T}$; the minimum eigenvalue varies from -20.716 to 6.852 and the maximum eigenvalue varies from -1.551 to 28.676. (d) Distribution of eigenvalues of the matrix $\mathbf{V}$ (the $\mathbf{S}$ matrix optimized as described in the article); the minimum eigenvalue varies from -0.001 to 2.815 and the maximum eigenvalue varies from 0 to 5.950.

Table 2
Examples of the TLS matrices.

The matrix elements extracted from the PDB files after refinement (§6).

PDB code	Chain, residue No.	$\mathbf{T}$ (Å^2)			$\mathbf{L}$ (deg^2)			$\mathbf{S}$ (Å deg)		
1dqv	A1–A97	0.1777	0.0090	−0.0044	1.4462	−0.0160	−0.2656	0.0467	−0.0523	0.0566
		0.0090	0.1306	0.0019	−0.0160	1.2556	0.4713	0.1010	0.0032	−0.0164
		−0.0044	0.0019	0.1372	−0.2656	0.4713	0.8689	0.0090	0.0188	0.0560
	B1–B97	0.1777	0.0090	−0.0044	1.4462	−0.0160	−0.2656	0.0467	−0.0523	0.0566
		0.0090	0.1306	0.0019	−0.0160	1.2556	0.4713	0.1010	0.0032	−0.0164
		−0.0044	0.0019	0.1372	−0.2656	0.4713	0.8689	0.0090	0.0188	0.0560
1exr	A2–A30	0.0899	0.0040	−0.0004	1.3491	−0.3760	−0.3971	−0.0249	−0.3537	−0.0874
		0.0040	0.1333	0.0058	−0.3760	0.6103	−0.3389	0.1275	0.0783	−0.0144
		−0.0004	0.0058	0.0728	−0.3971	−0.3389	0.3698	0.0183	0.0542	−0.0103
	A31–A74	0.0925	0.0037	0.0041	0.3464	0.3638	0.2923	−0.0220	−0.0419	−0.0793
		0.0037	0.0673	0.0062	0.3638	0.3283	0.1212	−0.0061	0.0018	0.1161
		0.0041	0.0062	0.1119	0.2923	0.1212	0.3799	−0.0041	−0.0385	−0.0009
	A75–A84	0.2433	0.0144	0.0917	0.0736	0.0171	0.0565	0.4357	0.1151	0.2346
		0.0144	0.2867	0.1720	0.0171	0.0068	−0.0203	−0.2521	−0.3549	−0.2041
		0.0917	0.1720	0.1749	0.0565	−0.0203	0.0336	−0.3793	−0.1499	0.0111
	A85–A147	0.0747	−0.0110	0.0066	0.6097	−0.0786	−0.1864	0.0180	0.1466	0.0378
		−0.0110	0.1384	0.0062	−0.0786	0.6474	−0.6233	0.0155	−0.0872	−0.0542
		0.0066	0.0062	0.0673	−0.1864	−0.6233	0.9637	−0.0440	0.1022	−0.0852
4b3x	A1–A65	0.4663	0.0991	−0.0764	0.4738	0.0063	0.2318	0.0391	−0.0307	−0.4316
		0.0991	0.5443	−0.0321	0.0063	0.2120	−0.0584	0.0587	0.1786	−0.2003
		−0.0764	−0.0321	0.5001	0.2318	−0.0584	0.1312	0.3665	0.4293	0.0403
	A66–A363	0.1649	−0.0259	0.0184	0.8808	−0.0912	−0.1736	−0.0345	0.0102	−0.0661
		−0.0259	0.1422	0.0055	−0.0912	0.9522	0.0972	0.1159	−0.0222	0.0999
		0.0184	0.0055	0.2028	−0.1736	0.0972	1.6563	0.0424	−0.1330	−0.0237

provides some examples of this procedure applied to models deposited in the PDB. §7 describes an example in which knowledge of the motion parameters extracted from the TLS matrices is necessary to explicitly simulate the ensemble of corresponding structures and the corresponding X-ray diffuse scattering.

6. Examples of TLS matrix analysis

As discussed in §1, there are numerous examples of fruitful application of the TLS formalism to structural studies of concerted motion. The goal of this section is to illustrate the algorithm described above, to describe possible traps that emerge during refinement and to discuss further developments.

6.1. Survey of available TLS matrices in the PDB

We have analyzed all available TLS matrices in the PDB. From an overall 106 761 entries (as of March 2015), 25 904 use TLS modelling. More than 20 000 of these entries have several TLS groups, resulting in a total of 203 261 sets of TLS matrices (Fig. 2a), with the largest number of groups per entry being 283 (PDB entry 3u8m; Rohde *et al.*, 2012). About a third of these sets have negative eigenvalues for the deposited $\mathbf{T}$ or $\mathbf{L}$ matrices. Some of these values are only slightly negative (Figs. 2b and 2c) and can be considered to be rounding errors, while the worst values are as small as −0.28 rad^2 for $\mathbf{L}$ and −20.72 Å^2 for $\mathbf{T}$. For 11 412 $\mathbf{T}$ matrices and 138 $\mathbf{L}$ matrices all three eigenvalues are negative.

Another third of the TLS groups cannot be interpreted by elemental motions owing to other reasons, as described in §§3 and 4 (Table 1).

After an initial screen to find the positive definite $\mathbf{T}$ and $\mathbf{L}$ matrices, we then ran a search for the elemental motions in two modes. Firstly, we tried to decompose the TLS matrices as taken directly from the PDB files. As expected, the average value of tr($\mathbf{S}$) is 3 × 10^{-5} Å (*i.e.* practically zero) and the corresponding r.m.s.d. is $\sigma = 10^{-2}$ Å. About 120 000 $\mathbf{S}$ matrices have |tr($\mathbf{S}$)| < 10^{-4} Å. The numbers of matrices with |tr($\mathbf{S}$)| larger than 1σ, 3σ, 10σ and 20σ are only 3772, 486, 31 and three, respectively. We then applied the aforementioned algorithm with the optimal choice of the value t_S to be subtracted from the diagonal $\mathbf{S}$ elements in each case.

Table 1 shows the results of both runs and illustrates that it is possible to fix the problems found in 6500 of the TLS sets (corresponding to about 500 PDB entries) by a correction of the diagonal elements of the $\mathbf{S}$ matrix as described above. The table takes into account possible rounding errors by correcting slightly negative eigenvalues (those closer in value to zero than 10^{-5} of the default units: Å^2, rad^2 and Å rad for $\mathbf{T}$, $\mathbf{L}$ and $\mathbf{S}$, respectively). For example, when running the algorithm in the $\mathbf{S}$ optimizing mode the program can formally calculate the $\mathbf{V}$ matrix for about 70 000 sets. For 2296 cases this matrix has negative eigenvalues (Fig. 2d), while in 2294 cases these eigenvalues are closer to 0 than $10^{-5}\,\text{Å}^2$; for such matrices the program makes automatic corrections and continues the process.

It is important to note that even if the parameters of the elemental motions can be formally extracted from the TLS matrices, this does not guarantee that they will make physical sense and therefore be valid for decomposition into a representative structural ensemble. Clearly, vibration amplitudes on the order of 20 Å^2 cannot represent harmonic vibrations (Fig. 2d). Similarly, the linear rotation approximation contained in TLS theory is valid only up to approximately 0.1 rad; however, much larger values can be found in the PDB (Fig. 2b). Similar restrictions also hold for the screw parameters. The products $s_x d_x, s_y d_y, s_z d_z$ show the mean shifts along the screw axes owing to librations around these axes; the values found in the PDB approaching 3 Å seem to be too large to describe harmonic motions.

For a more detailed analysis, we selected several entries from the PDB. For each structure, we applied a standard TLS refinement protocol as implemented in *phenix.refine* (Afonine *et al.*, 2012) including automatic determination of the TLS groups. During refinement, 20 matrix elements were refined independently, six for $\mathbf{T}$, six for $\mathbf{L}$ and eight for $\mathbf{S}$; the three

Table 3
Examples of parameters of the elemental motions found from decomposition of the TLS matrices.

The parameters are given in the units used in this article, allowing an easy estimation of the corresponding atomic displacements. The directions of the libration and rotation axes are not given.

| PDB code	Chain, residue No.	**T**: t_x, t_y, t_z (Å)	**L**: d_x, d_y, d_z (rad)	**S**: s_x, s_y, s_z (Å)	tr(**S**)
1dqv	A1–A97	0.3455 0.3671 0.4172	0.01239 0.02044 0.02273	1.343 1.137 −1.319	0
	B1–B97	0.3634 0.3885 0.4166	0.01608 0.01753 0.03069	0.679 −1.177 0.200	0
1exr	A2–A30	0.1944 0.2663 0.2870	0.00000 0.01602 0.02182	0.000 2.951 3.408	>0
	A31–A74	0.2110 0.2939 0.3068	0.00000 0.00860 0.01637	0.000 −18.14 −5.028	<0
	A75–A84	0.1692 0.4906 0.6598	0.00000 0.00000 0.00000	0.000 0.000 0.000	0
	A85–A147	0.0002 0.2270 0 3078	0.00553 0.01418 0.02109	20.83 0.800 −1.672	∼0
4b3x	A1–A65	0.0994 0.6064 0.7116	0.00000 0.00825 0.01343	0.000 2.718 −11.05	<0
	A66–A363	0.3306 0.4102 0.4413	0.01568 0.01720 0.02283	3.164 −2.276 −0.197	0

(a)

(b)

(c)
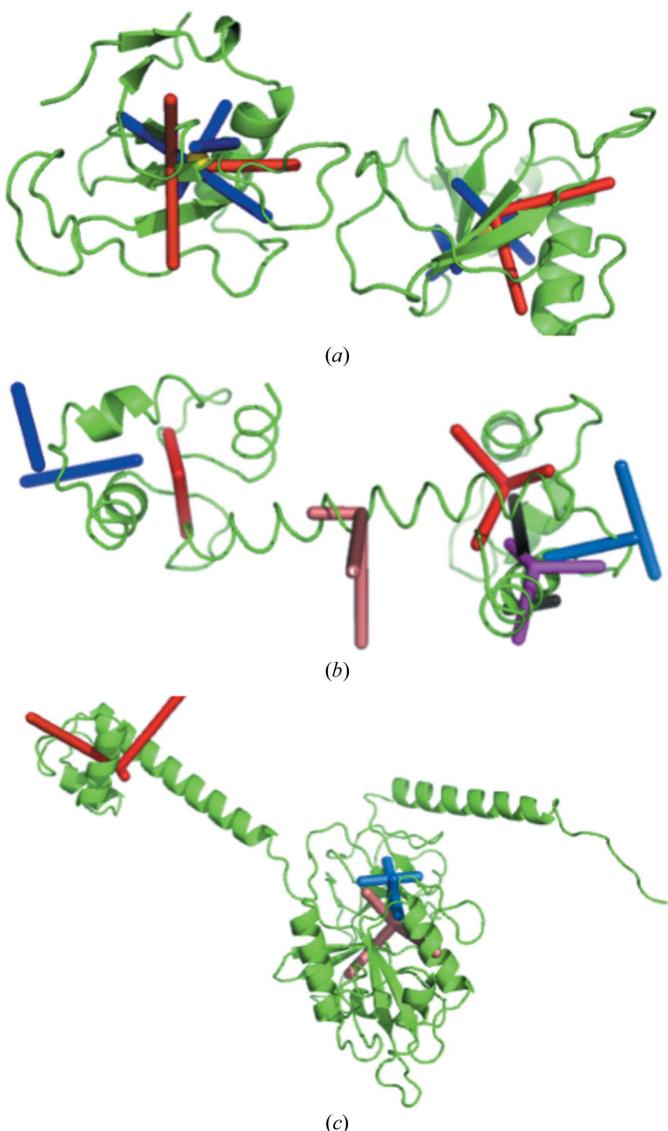
Figure 3
Examples of the vibration–libration ensembles. Red/salmon/magenta sticks indicate the principal vibration axes, with the origin in the centre of the group; blue/marine/black sticks are the libration axes. Yellow spheres in the 1dqv model show the reaction centres. (a) 1dqv model. (b) 1exr model; note pure vibrations for group 3 (the helix) and the absence of one of the libration axes for groups 1 and 2. (c) 4b3x model. Libration axes for the first group are not shown as they are too far from the molecule.

diagonal elements of **S** were constrained such that the trace of the matrix is equal to 0. The procedure described above (§§3–5) was then applied to all sets of obtained TLS matrices.

We remind the reader that the elements of the **L** and **S** matrices are expressed in rad^2 and Å rad, while in the PDB files they are in deg^2 and in Å deg, respectively (Table 2).

6.2. Synaptotagmin

The crystals of synaptotagmin III (PDB entry 1dqv; Sutton *et al.*, 1999) contain two copies of the molecule in the asymmetric unit. The structure after re-refinement by *phenix.refine* without TLS modelling has an R_{work} of 0.200 and an R_{free} of 0.231 at a resolution of 2.5 Å. Performing TLS refinement with each molecule taken as a single TLS group reduced the R factors to $R_{work} = 0.177$ and $R_{free} = 0.211$, indicating that this additional modelling significantly improves the agreement with the experimental data. Table 2 shows the two sets of matrices and Table 3 contains the corresponding motion parameters extracted using our approach. For the two groups both vibrations and librations are practically isotropic and are of the same order of magnitude. Fig. 3(a) shows the principal axes of these motions.

6.3. Calmodulin

The structure of calmodulin (PDB entry 1exr; Wilson & Brunger, 2000) has been determined previously at a resolution of 1.0 Å. This example illustrates possible problems that can be solved by a minimal correction of the TLS values. For re-refinement with *phenix.refine* the model was automatically split into four TLS groups. For the first group, one of the eigenvalues of the matrix **L** was equal to -2×10^{-5} rad^2. If we consider this value to be zero (in this case the zero value must be also assigned to off-diagonal elements of the first row of the matrix **S**), the composite motions contain only two libration axes and their parameters can be extracted. Corresponding modifications of the resulting matrices $\mathbf{U}_{group,n}$ (2) can be compensated for by respective adjustment of the individual contributions $\mathbf{U}_{local,n}$. This keeps the total ADP parameters $\mathbf{U}_{Cart,n}$ unchanged, thus maintaining the previously calculated structure factors and R factors. An accurate separation of total atomic displacement parameter values into contributions from several sources (see, for example, Murshudov *et al.*, 1999; Winn *et al.*, 2001, 2003; Afonine *et al.*, 2012) is a separate ongoing project (Afonine & Urzhumtsev, 2007).

For the second TLS group, the refined TLS matrix elements contained one degenerate libration. The procedure described in §§3–5 was successfully applied. Note that this procedure modified the diagonal elements of the matrix **S**, removing an appropriate value of the parameter t_S (§4.4) and making tr(**S**) nonzero.

For the third group, all three eigenvalues of the matrix **L** were extremely small (0.0, 0.8×10^{-5} and 3×10^{-5} rad^2), producing high computational instability and extremely large screw parameters that resulted in the inability of the procedure to find a positive semidefinite $\mathbf{V}_{[L]}$ (27). If we define all librations to be absent and replace matrix **L** (and respectively **S**) by zero matrices, the vibration parameters can easily be found from **T**. In fact, this TLS group is a helix held at both ends by large domains, which leads to the expectation of a pure vibration motion.

Finally, for the fourth group one of the diagonal elements of the matrix **T** was marginally negative. Increasing all of the diagonal elements of the matrix **T** by 0.002 Å^2 makes this matrix positive definite (this corresponds to $B = 0.16$ Å^2). As discussed above, this adjustment can be compensated for by removing the equivalent amount from individual atomic contributions $\mathbf{U}_{\mathrm{local},n}$ (such a subtraction keeps the individual atomic contributions positive). This group vibrates in a plane (Fig. 3b) and the principal vibration axis of group 3 (the helix) is parallel to this plane, leading to the plausible hypothesis that groups 3 and 4 at least partially move together or slide along each other.

To check the influence of the manual modification on the TLS matrices, we recalculated the R factors before and after performing these changes without updating the individual atomic contributions $\mathbf{U}_{\mathrm{local},n}$. For all of the modifications described above, including the ensemble of modifications applied together, the R factors only varied in the fourth significant digit.

This example demonstrates that although current refinement procedures may result in TLS matrices that are unable to satisfy the previously mentioned conditions, small changes to them may provide sufficient correction. This highlights the need to use appropriate restraints or constraints on refinable parameters within the TLS model.

6.4. Initiation translation factor 2 (IF2)

The structure of IF2 (PDB entry 4b3x) has recently been solved in one of our laboratories (Simonetti *et al.*, 2013) with an R_{work} of 0.180 and an R_{free} of 0.219 at a resolution of 1.95 Å. *A posteriori* TLS refinement was performed with two groups: the first group included the N-terminus and the following long helix, and the second included the rest of structure. Re-refining the model produced better statistics, with $R_{\mathrm{work}} = 0.176$ and $R_{\mathrm{free}} = 0.203$. In this example, the TLS matrices from the first group were not directly interpretable because the residual matrix $\mathbf{V}_{[L]}$ was not positive semidefinite (the minimal eigenvalue was -0.05). Similarly to the last TLS group in calmodulin, we artificially added 0.06 Å^2 to all diagonal elements of the matrix **T**, corresponding to roughly 5 Å^2 (the same amount has been removed from the residual atomic B values, thus leaving the R factors unchanged). This correction allowed interpretation of the TLS matrices in terms of elemental motions. We note that for the first TLS group one of the rotations was degenerate and that the assignment tr(**S**) = 0 would make this matrix incompatible with **L**. Table 3 shows

that the vibrations of this group are essentially anisotropic. Fig. 3(c) also shows that the libration axes for this group pass quite far away from the molecule, which makes the corresponding rotation similar to a translation. Additionally, we believe that the large s_z value indicates that the matrix **S** is not well defined. The matrices for the second group were interpreted and revealed isotropic vibrations and librations.

Finally, we tried to apply the same procedure after choosing the TLS groups manually as residues 1–50 (N-terminus), 51–69 (helix), 70–333 (G domain) and 343–363 (the connector to the C domain, which is absent in this structure). As before, the matrices were interpretable for the G domain. For groups 2 and 4, after an adjustment similar to those discussed above (a slight increase of the diagonal **T** elements with a decrease of the residual atomic B values), we obtained a pure vibration for the helix (as for the calmodulin case) and a libration around a single axis for the terminal group. In contrast, we failed to find reasonably small corrections for the matrices of the first group that would make them interpretable in terms of physical motions that in particular can be represented by a structural ensemble.

This case exemplifies a situation in which the current TLS refinement protocols result in matrices that significantly reduce the R factors without providing refined TLS parameters that can be decomposed into a physically realistic motion of one of the groups. This highlights the need to improve TLS refinement algorithms by making use of constraints on aforementioned conditions on TLS matrices.

7. Interpreting TLS matrices with a structural ensemble

7.1. Generation of an explicit set of atomic models with a variability consistent with TLS

Some structural problems may explicitly require a set of models that describe a given mobility, for example corresponding to the TLS matrices for harmonic motion. An example of such a problem is described in the accompanying paper by Van Benschoten *et al.* (2015) (and is briefly presented in §7.4), in which X-ray diffuse scattering data were compared with calculated data corresponding to different types of molecular motion. Other examples may include analyzing larger-scale anharmonic motions, for which techniques such as molecular-dynamics trajectories have traditionally been used (McCammon *et al.*, 1977).

If a model deposited in the PDB contains TLS matrices, the matrices can be decomposed as described above. As soon as a combination of vibrations and librations is extracted from the TLS matrices, we can explicitly build a corresponding set of models. Knowledge of the three vibrations and three librations provides the atomic shifts underlying the total displacement.

It is generally more convenient to generate each group of atomic shifts in its own basis: shifts $\Delta_{[V]}^V \mathbf{r}_n$ owing to vibration in the [V] basis and shifts $\Delta_{[L]}^L \mathbf{r}_n$ owing to libration in the [L] basis. Here, we are working in a linear approximation such that rotation angles are on the order of 0.1 rad. For each

particular set of generated shifts, they are transformed into the [M] basis as $\Delta_{[M]}^{V}\mathbf{r}_n$ and $\Delta_{[M]}^{L}\mathbf{r}_n$ and their sum,

$$\Delta_{[M]}\mathbf{r}_n = \Delta_{[M]}^{L}\mathbf{r}_n + \Delta_{[M]}^{V}\mathbf{r}_n, \qquad (42)$$

is applied to the corresponding atoms. Details of model generation are discussed in the next sections. This procedure is repeated independently multiple times, leading to structural models distributed according to the TLS matrices.

7.2. Calculation of the model shift owing to libration

Let us suppose that we know (in the basis [M]) the direction of the three mutually orthogonal axes $\mathbf{l}_x$, $\mathbf{l}_y$, $\mathbf{l}_z$ for independent libration as well as the coordinates of the points $\mathbf{w}_{[M]}^{lx}$, $\mathbf{w}_{[M]}^{ly}$, $\mathbf{w}_{[M]}^{lz}$ belonging to each axis. We recalculate the coordinates of these points and the coordinates $(x_{[M]n}, y_{[M]n}, z_{[M]n})$, $n = 1, 2, \ldots, N$, of all atoms $\mathbf{r}_{[M]n}$ of the group into the [L] basis as

$$\mathbf{r}_{[L]n} = \mathbf{R}_{ML}^{-1}\mathbf{r}_{[M]n} = \mathbf{R}_{ML}^{\tau}\mathbf{r}_{[M]n} \qquad (43)$$

(similar relations are derived for the points $\mathbf{w}_{[M]}^{lx}$, $\mathbf{w}_{[M]}^{ly}$, $\mathbf{w}_{[M]}^{lz}$). We remind the reader that the squared libration amplitudes $\langle d_x^2 \rangle = L_{[L]xx} = \lambda_1$, $\langle d_y^2 \rangle = L_{[L]yy} = \lambda_2$, $\langle d_z^2 \rangle = L_{[L]zz} = \lambda_3$ (§3.2) and the screw parameters s_x, s_y, s_z (§4.5) are independent of the basis.

For an atom at a distance $R = 1$ Å from the rotation axis, the probability of the shifts d_x, d_y, d_z, which are numerically equal to the rotation angle in radians, are equal to

axis parallel to $\mathbf{l}_x$: $P(d_x) = (2\pi\lambda_1)^{1/2}\exp(-d_x^2/2\lambda_1)$,

axis parallel to $\mathbf{l}_y$: $P(d_y) = (2\pi\lambda_2)^{1/2}\exp(-d_y^2/2\lambda_2)$,

axis parallel to $\mathbf{l}_z$: $P(d_z) = (2\pi\lambda_3)^{1/2}\exp(-d_z^2/2\lambda_3)$. $\qquad (44)$

If one of the eigenvalues is equal to 0 then the corresponding d is equal to 0 with unit probability. The particular values of d_{x0}, d_{y0}, d_{z0} are obtained using a random-number generator with an underlying normal distribution (44).

For each of the axes $\mathbf{l}_x$, $\mathbf{l}_y$, $\mathbf{l}_z$ for each atom n described by the vector $\mathbf{r}_n$, we calculate the coordinates, in the [L] basis, of its shifts $\Delta_{[L]}^{lx}\mathbf{r}_n$, $\Delta_{[L]}^{ly}\mathbf{r}_n$, $\Delta_{[L]}^{lz}\mathbf{r}_n$ owing to the corresponding rotations by d_{x0}, d_{y0}, d_{z0} (Appendix A). The overall shift owing to libration around the three axes is the sum

$$\Delta_{[L]}^{L}\mathbf{r}_n = \Delta_{[L]}^{lx}\mathbf{r}_n + \Delta_{[L]}^{ly}\mathbf{r}_n + \Delta_{[L]}^{lz}\mathbf{r}_n. \qquad (45)$$

It changes from one atom of the group to another and must be calculated for all atoms of the group with the same (d_{x0}, d_{y0}, d_{z0}) values for a particular instance of the three rotations.

To transform the atomic shift (45) from the [L] basis into the initial [M] basis, we invert (43),

$$\Delta_{[M]}^{L}\mathbf{r}_n = \mathbf{R}_{ML}\Delta_{[L]}^{L}\mathbf{r}_n. \qquad (46)$$

7.3. Calculation of the model shift owing to vibration

In the harmonic approximation, the independent vibration shifts t_x, t_y, t_z expressed in the [V] basis are distributed accordingly to the probability laws

**Entire molecule**　　**Monomer**　　**Subdomain**
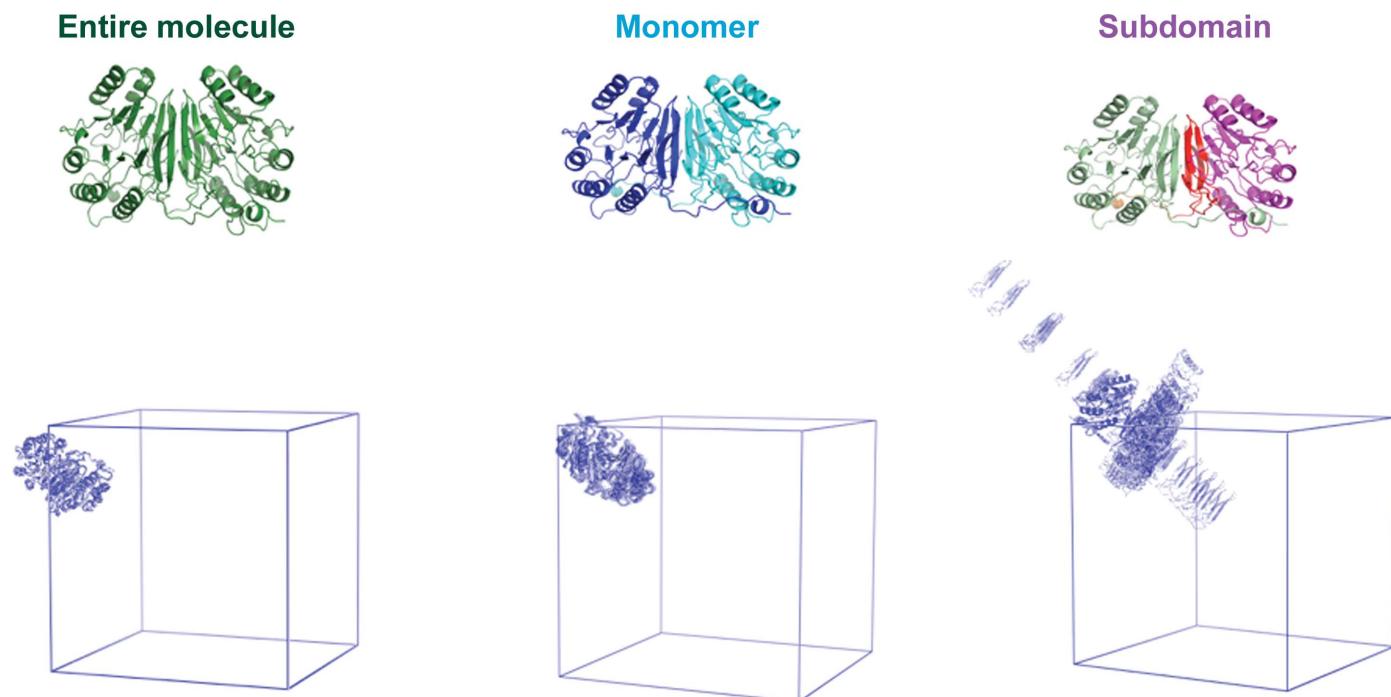
Figure 4
GpdQ TLS ensembles. The GpdQ TLS groups are projected onto the protein structure. The corresponding ensembles produced by *phenix.tls_as_xyz* are shown below. Each TLS PDB ensemble is shown as a single asymmetric unit outlined by the unit cell. An increase in overall motion is apparent going from left to right. The 20-member ensemble is shown for visual simplicity.

$$P(t_x) = (2\pi V_{[V]xx})^{1/2} \exp(-t_x^2/2V_{[V]xx}) = (2\pi\mu_1)^{1/2} \exp(-t_x^2/2\mu_1),$$

$$P(t_y) = (2\pi V_{[V]yy})^{1/2} \exp(-t_y^2/2V_{[V]yy}) = (2\pi\mu_2)^{1/2} \exp(-t_y^2/2\mu_2),$$

$$P(t_z) = (2\pi V_{[V]zz})^{1/2} \exp(-t_z^2/2V_{[V]zz}) = (2\pi\mu_3)^{1/2} \exp(-t_z^2/2\mu_3).$$

$$(47)$$

Using a random-number generator, for each model we obtain particular values of t_{x0}, t_{y0}, t_{z0} using (47). If one of the eigenvalues μ is equal to zero, the zero value is assigned to the corresponding shift. The overall translational shift, common to all atoms of the rigid group, is equal to

$$\Delta_{[V]}^{V}\mathbf{r}_n = t_{x0}\mathbf{v}_x + t_{y0}\mathbf{v}_y + t_{z0}\mathbf{v}_z. \tag{48}$$

In order to obtain this shift in the [M] basis, we calculate, similarly to (46),

$$\Delta_{[M]}^{V}\mathbf{r}_n = \mathbf{R}_{MV}\Delta_{[V]}^{V}\mathbf{r}_n. \tag{49}$$

7.4. Validation and application to GpdQ

We generated the ensembles produced by alternative TLS refinements of the glycerophosphodiesterase GpdQ (Jackson *et al.*, 2007). GpdQ is found in *Enterobacter aerogenes* and contributes to the homeostasis of the cell membrane by hydrolyzing the $3'$–$5'$ phosphodiester bond in glycerophos-

phodiesters. Each dimer contains three distinct domains per monomer: an α/β sandwich fold containing the active site, a domain-swapped active-site cap and a novel dimerization domain comprised of dual-stranded antiparallel β-sheets connected by a small β-sheet. Owing to the high global B factors and the presence of diffuse signal (Fig. 4), Jackson *et al.* (2007) performed three separate TLS refinements to model the crystalline disorder: entire molecule, monomer and subdomain. All TLS refinement attempts improved the R_{free} values when compared with the standard isotropic B-factor refinement; however, there was no significant difference among the final R_{free} values from the various TLS runs. We hypothesized that the diffuse scattering produced by each TLS motion would contain significant differences, as diffuse signal is a direct result of correlated motion. The notion that TLS refinement produces unique diffuse signal has been suggested previously (Tickle & Moss, 1999). Physical ensembles of the TLS motion, rather than a mathematical description, were required to generate three-dimensional diffuse scattering maps from *phenix.diffuse*. Visual inspection confirmed that the ensembles produced by *phenix.tls_as_xyz* replicated the anisotropic motion predicted by TLS thermal ellipsoids (Fig. 5). Additionally, we calculated the structure factors predicted by the original TLS refinement 'entire molecule' and compared them with the F_{model} values (for example, as
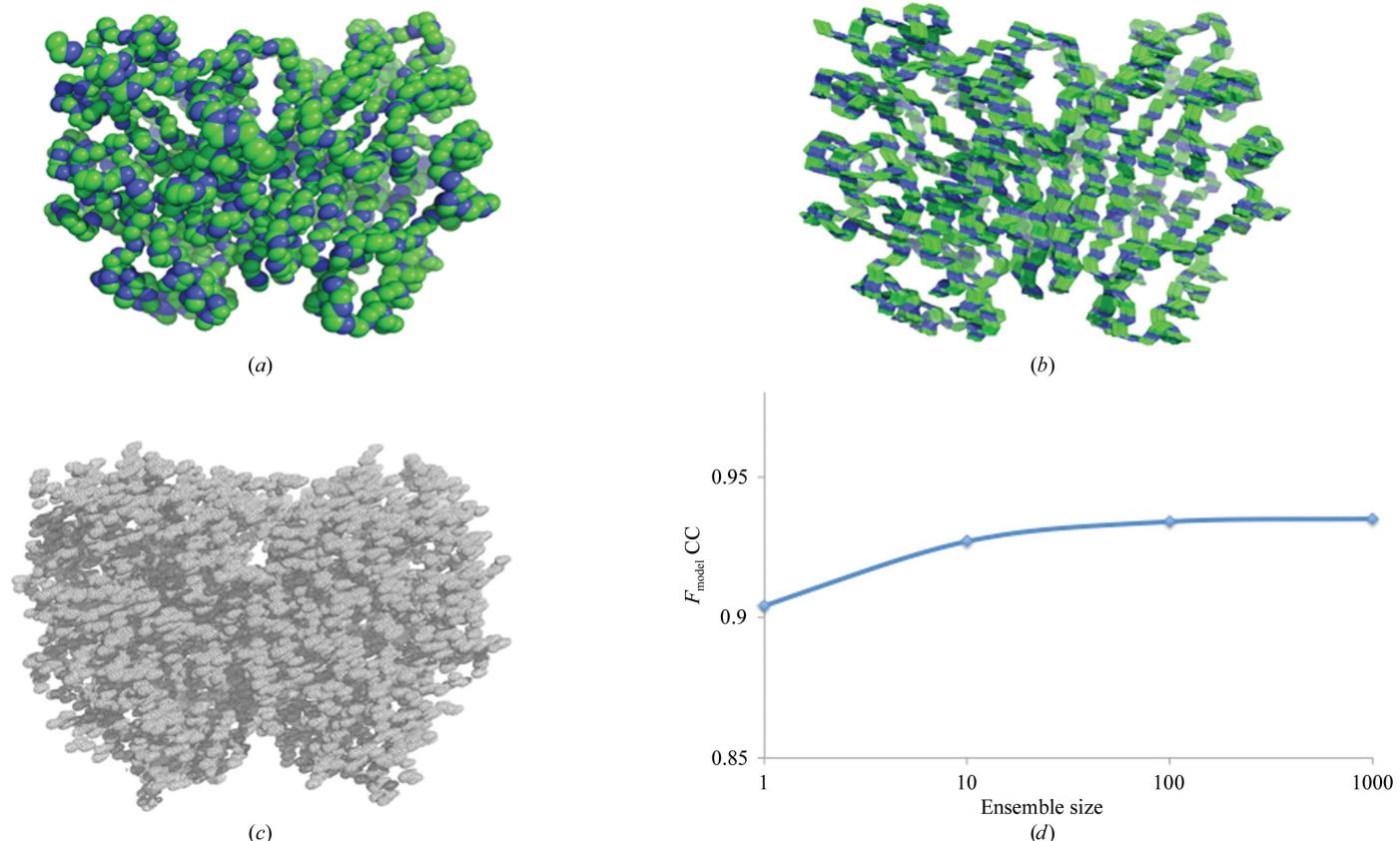

Figure 5

phenix.tls_as_xyz ensembles replicate TLS anisotropic motion. (*a*) GpdQ backbone with thermal ellipsoid representation of 'entire molecule' TLS anisotropic B factors. (*b*) *phenix.tls_as_xyz* ensemble backbones produced from 'entire molecule' TLS refinement. (*c*) Complete electron density predicted by 'entire molecule' TLS refinement. (*d*) Global correlation coefficient between experimental structure-factor amplitudes F_{obs} of the original GpdQ 'entire motion' refinement and *phenix.tls_as_xyz* ensembles of various sizes. Convergence values plateau at 0.935.

defined in Afonine *et al.*, 2012) produced by various *phenix.tls_as_xyz* ensemble sizes. The structure factors converged to a global correlation value of 0.965, demonstrating that *phenix.tls_as_xyz* ensembles accurately represent the motions predicted by TLS refinement. Physical representation of the underlying motion also revealed that while two of the TLS refinements produced motion with small variances (a necessity within TLS theory), using each functional region as a TLS group produced fluctuations that are clearly improbable (Fig. 4). Thus, viewing TLS refinement in the form of a structural ensemble is a valuable check of the validity of the results, as matrix elements that satisfy the previously described conditions may still produce motions that are clearly implausible.

8. Discussion

While our previous review on the subject (Urzhumtsev *et al.*, 2013) described the computational details of obtaining the TLS matrices from a known set of vibration and libration parameters (including the position of the axes and correlation of these motions), the current work focuses on the opposite problem of extracting these parameters from a given set of TLS matrices. The problem is not as simple as it may at first seem.

This difficulty arises because current structure-refinement programs vary the matrix elements as independent parameters and often ignore critical constraints on real-space motions. A second challenge is that identical motions may be represented by different vibration–libration combinations. As a consequence, there is no one-to-one relationship between these parameters and the set of TLS matrices. In particular, the traditional way of choosing the matrix **S** so that its trace is equal to zero may result in a mutually inconsistent combination of TLS matrices.

This manuscript describes the constraints that can be used to validate a given set of **T**, **L** and **S** matrices and to improve the refinement of TLS parameters. Beyond the well known conditions of non-negativity for the eigenvalues of **T** and **L**, we also discuss the conditions that relate the matrices, a crucial step in ensuring that the results of TLS refinement correspond to physically possible combinations of librations and vibrations. Taking all these conditions into account provides the possibility of correcting TLS matrices in some cases, if needed. Building these conditions into refinement protocols can eliminate nonplausible refined TLS matrices

The TLS matrix representation is a convenient way of encoding concerted motions into a form suitable for the calculation of structure factors and, in turn, structure refinement. There are two drawbacks to the standard implementation of this method. Firstly, TLS matrices cannot readily be interpreted in terms of underlying motions, but rather require additional processing in order for this information to be extracted. Secondly, direct refinement of the TLS matrix elements may result in refined matrices that cannot be represented as a structural ensemble. To address these two drawbacks, we propose using the set of vibration and libration

parameters as refinable variables (an ongoing project for the authors) and reporting them in the PDB files. Indeed, using actual motion descriptors as refinement variables will allow more effective application of physical constraints and in turn guarantee that refined values can be translated to structural ensembles, simplifying the analysis of refinement results, as they will be readily available for interpretation. Finally, this strategy will reduce the chance of overfitting data with atomic models that represent implausible concerted motions.

The survey of PDB entries refined with TLS revealed that roughly 85% of these deposited models contain matrices that are not consistent with the underlying physical model of the concerted motions. Therefore, these matrices cannot be interpreted in terms of rigid-body rotations and translations, and in turn cannot represent these motions (Table 1). This highlights two urgent needs. Firstly, existing refinement programs should be updated so that they apply appropriate restraints or constraints on refinable parameters of the TLS model. This should be followed by the implementation and use of comprehensive validation of TLS refinement results.

The utility of our presented algorithm is twofold: it validates TLS matrices to confirm that they can represent concerted structural motions and interprets TLS matrices in terms of the elemental motions that they describe. The information about atomic group motions conveyed by the TLS model can be used to analyze possible molecular mechanisms (as illustrated previously). Descriptions of TLS motion may also be used to generate an ensemble of molecular conformations, from which the predicted diffuse scattering signal can be calculated (see the accompanying paper by Van Benschoten *et al.*, 2015.).

The current procedures for analyzing and validating TLS parameters, as well as the algorithm for generating a set of models from given libration and vibration parameters, are implemented in the *PHENIX* suite and are called *phenix. tls_analysis* and *phenix.tls_as_xyz*, respectively. The programs are available starting with version dev-1890.

APPENDIX *A*
Technical details of the algorithm

A1. Definition of the transition matrices

Let us have three mutually orthogonal unit vectors $\mathbf{l}_x$, $\mathbf{l}_y$, $\mathbf{l}_z$ described respectively by their coordinates $[(l_x)_{[M]x}, (l_x)_{[M]y}, (l_x)_{[M]z}]$, $[(l_y)_{[M]x}, (l_y)_{[M]y}, (l_y)_{[M]z}]$, $[(l_z)_{[M]x}, (l_z)_{[M]y}, (l_z)_{[M]z}]$ in the Cartesian basis [M]. These vectors can be considered as a new basis [L]. The coordinates of a vector $\mathbf{r}$ in [L] and [M] are expressed through each other using the transition matrix $\mathbf{R}_{ML}$ as

$$\begin{pmatrix} x_{[M]} \\ y_{[M]} \\ z_{[M]} \end{pmatrix} = \mathbf{R}_{ML} \begin{pmatrix} x_{[L]} \\ y_{[L]} \\ z_{[L]} \end{pmatrix} = \begin{bmatrix} (l_x)_{[M]x} & (l_y)_{[M]x} & (l_z)_{[M]x} \\ (l_x)_{[M]y} & (l_y)_{[M]y} & (l_z)_{[M]y} \\ (l_x)_{[M]z} & (l_y)_{[M]z} & (l_z)_{[M]z} \end{bmatrix} \begin{pmatrix} x_{[L]} \\ y_{[L]} \\ z_{[L]} \end{pmatrix}. \quad (50)$$

Transition matrices for other pairs of bases, for example from [V] to [L] (§2.1), from [M] to [V] and *vice versa* (§7.3) are defined in a similar way.

A2. Cauchy conditions on the elements of the TLS matrices

Let d_x, d_y, d_z and u_x, u_y, u_z be random displacements owing to rotations and translations, respectively. Since $S_{xx} = \langle d_x u_x \rangle$, $V_{xx} = \langle u_x u_x \rangle$, $L_{xx} = \langle d_x d_x \rangle$ (Schomaker & Trueblood, 1968; see also equations 8.5–8.7 in Urzhumtsev et al., 2013), it follows from the Cauchy inequality that

$$S_{xx}^2 \leq T_{xx} L_{xx}. \tag{51}$$

In the basis [L] with $\mathbf{S}_{[L]} = \mathbf{S}_C(t_S)$ (18), condition (51) becomes

$$(S_{[L]xx} - t)^2 \leq T_{C[L]xx} L_{[L]xx}. \tag{52}$$

Similarly, we obtain two other conditions

$$(S_{[L]yy} - t)^2 \leq T_{C[L]yy} L_{[L]yy}, \quad (S_{[L]zz} - t)^2 \leq T_{C[L]zz} L_{[L]zz}. \tag{53}$$

A3. Polynomials for the coefficients of the characteristic equation

If t_{xx}, t_{xy}, t_{xz} etc. are respective elements of the matrix $\mathbf{T}_\Lambda$ (29), the coefficients (36) of the characteristic equation as functions of the parameter t are

$$a_S(t) = [(t - S_{[L]xx})^2 - t_{xx}] + [(t - S_{[L]yy})^2 - t_{yy}] + [(t - S_{[L]zz})^2 - t_{zz}], \tag{54}$$

$$\begin{aligned} b_S(t) = &[(t - S_{[L]xx})^2 - t_{xx}][(t - S_{[L]yy})^2 - t_{yy}] \\ &+ [(t - S_{[L]yy})^2 - t_{yy}][(t - S_{[L]zz})^2 - t_{zz}] \\ &+ [(t - S_{[L]zz})^2 - t_{zz}][(t - S_{[L]xx})^2 - t_{xx}] \\ &- (t_{xy}^2 + t_{yz}^2 + t_{zx}^2), \end{aligned} \tag{55}$$

$$\begin{aligned} c_S(t) = &[(t - S_{[L]xx})^2 - t_{xx}][(t - S_{[L]yy})^2 - t_{yy}][(t - S_{[L]zz})^2 - t_{zz}] \\ &- t_{yz}^2[(t - S_{[L]xx})^2 - t_{xx}] - t_{xz}^2[(t - S_{[L]yy})^2 - t_{yy}] \\ &- t_{xy}^2[(t - S_{[L]zz})^2 - t_{zz}] - 2t_{xy}t_{yz}t_{xz}. \end{aligned} \tag{56}$$

A4. Explicit expression for the atomic shifts owing to rotations with given parameters

Let $(x_{[L]}, y_{[L]}, z_{[L]})$ be the Cartesian coordinates of a point $\mathbf{r}$ in the basis [L]. For a rotation around the axis parallel to $\mathbf{l}_z$ and crossing the point $\mathbf{w}_{[L]}^{lz} = (w_{[L]x}^{lz}, w_{[L]y}^{lz}, w_{[L]z}^{lz})$, we first recalculate the coordinates of the vector $\mathbf{r} - \mathbf{w}_{[L]}^{lz}$ with respect to the rotation axis,

$$x_{[A]} = x_{[L]} - w_x^{lz}; \ y_{[A]} = y_{[L]} - w_y^{lz}; \ z_{[A]} = z_{[L]n} - w_z^{lz}. \tag{57}$$

If $\mathbf{r}'$ stands for the position of the same point after rotation by angle d_{z0} around this axis, the coordinates of $\mathbf{r}' - \mathbf{w}_{[L]}^{lz}$, the point with respect to the axis, are

$$(x_{[A]} \cos d_{z0} - y_{[A]} \sin d_{z0}); \ (x_{[A]} \sin d_{z0} + y_{[A]} \cos d_{z0}); \\ (z_{[A]} + s_z d_{z0}). \tag{58}$$

This gives the atomic shift

$$\begin{aligned} \Delta_{[L]}^{lz} \mathbf{r} = \mathbf{r}' - \mathbf{r} &= (\mathbf{r}' - \mathbf{w}_{[L]}^{lz}) - (\mathbf{r} - \mathbf{w}_{[L]}^{lz}) \\ &= [(x_{[L]} - w_x^{lz})(\cos d_{z0} - 1) - (y_{[L]} - w_y^{lz}) \sin d_{z0}]\mathbf{l}_x \\ &\quad + [(x_{[L]} - w_x^{lz}) \sin d_{z0} + (y_{[L]} - w_y^{lz})(\cos d_{z0} - 1)]\mathbf{l}_y \\ &\quad + s_z d_{z0} \mathbf{l}_z. \end{aligned} \tag{59}$$

There are similar expressions for the shift owing to rotations around the other two axes:

$$\begin{aligned} \Delta_{[L]}^{lx} \mathbf{r} &= [(y_{[L]} - w_y^{lx})(\cos d_{x0} - 1) - (z_{[L]} - w_z^{lx}) \sin d_{x0}]\mathbf{l}_y \\ &\quad + [(y_{[L]} - w_y^{lx}) \sin d_{x0} + (z_{[L]} - w_z^{lx})(\cos d_{x0} - 1)]\mathbf{l}_z \\ &\quad + s_x d_{x0} \mathbf{l}_x, \end{aligned} \tag{60}$$

$$\begin{aligned} \Delta_{[L]}^{ly} \mathbf{r} &= [(z_{[L]} - w_z^{ly})(\cos d_{y0} - 1) - (x_{[L]} - w_x^{ly}) \sin d_{y0}]\mathbf{l}_z \\ &\quad + [(z_{[L]} - w_z^{ly}) \sin d_{y0} + (x_{[L]} - w_x^{ly})(\cos d_{y0} - 1)]\mathbf{l}_x \\ &\quad + s_y d_{y0} \mathbf{l}_y. \end{aligned} \tag{61}$$

APPENDIX B
List of abnormal situations requiring interruption of the procedure

This appendix summarizes the situations when the described algorithm breaks. Each condition below starts from the corresponding program message and then refers to the main text and to Fig. 1. To analyze the PDB content, the program can be run in a special regime when at step C we assign $t_S = 0$, i.e. when the matrix $\mathbf{S}$ is taken without any correction [in most cases this corresponds to the current default constraint tr($\mathbf{S}$) = 0]. In this regime, we directly calculate the matrices $\mathbf{C}$ and check the conditions (x)–(xii).

Step A: basis [L]; determination of the libration axes and amplitudes.

(i) 'Input matrix L[M] is not positive semidefinite', §3.1.

(ii) 'Input matrix T[M] is not positive semidefinite', §3.1.

Step B: determination of the points $\mathbf{w}$ at the libration axes.

(iii) 'Non-zero off-diagonal S[L] and zero L[L] elements', §3.2, (15).

(iv) 'Matrix T_C[L] is not positive semidefinite', §3.2, (17).

Step C: determination of the screw parameters: left branch (librations around all three axes).

(v) 'Empty (tmin_c, tmax_c) interval', §4.2, (23). $t_{\min,C} > t_{\max,C}$.

(vi) 'Empty (tmin_t, tmax_t) interval', §4.3, (31). $t_{\min,\tau} > t_{\max,\tau}$.

(vii) 'Negative argument when estimating tmin_a', §4.3, (38).

(viii) 'Intersection of the intervals for t_S is empty', §4.3, step (ii). $t_{\min} > t_{\max}$.

(ix) 't_min = t_max giving non positive semidefinite V_lambda', §4.3, step (iii).

(x) 'Interval (t_min, t_max) has no t value giving positive semidefinite V', §4.3, step (iv).

Step C: determination of the screw parameters: right branch (no libration around at least one of the axes).

(xi) 'Cauchy-Schwarz conditions are wrong for the found t_S', (22) with t_S calculated in §4.4.

(xii) 'Non-zero diagonal S[L] element for a zero L[L] element', §4.4.

Step D: determination of the vibration parameters.

(xiv) 'Matrix V[L] is not positive semidefinite', §5.1.

Extra checks at step C when some conditions may fail owing to rounding errors.

(1) When calculating square roots in (24), the arguments are non-negative by previous conditions (i) and (iv) since the diagonal elements of a positive semidefinite matrix are non-negative.

(2) When calculating square roots in (28), the arguments are non-negative by previous condition (i).

(3) When calculating square roots in (31), the argument τ_{max} is non-negative since the eigenvalues of $\mathbf{T}_{C[L]}$ are also non-negative by previous condition (iv).

Acknowledgements

References

Adams, P. D. et al. (2010). Acta Cryst. D**66**, 213–221.

Afonine, P. V., Grosse-Kunstleve, R. W., Adams, P. D. & Urzhumtsev, A. (2013). Acta Cryst. D**69**, 625–634.

Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H. & Adams, P. D. (2012). Acta Cryst. D**68**, 352–367.

Afonine, P. & Urzhumtsev, A. (2007). CCP4 Newsl. Protein Crystallogr. **45**, contribution 6.

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.

Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). J. Mol. Biol. **112**, 535–542.

Brenner, H. (1967). J. Colloid Interface Chem. **23**, 407–435.

Chaudhry, C., Horwich, A. L., Brunger, A. T. & Adams, P. D. (2004). J. Mol. Biol. **342**, 229–245.

Coppens, P. (2006). International Tables for Crystallography, Vol. B, edited by U. Shmueli, pp. 10–24. Dordrecht: Kluwer Academic Publishers.

Cruickshank, D. W. J. (1956). Acta Cryst. **9**, 754–756.

Dauter, Z., Murshudov, G. N. & Wilson, K. S. (2012). International Tables for Crystallography, Vol. F, 2nd ed., edited by E. Arnold, D.

M. Himmel & M. G. Rossmann, pp. 485–498. Chichester: Wiley.

Dunitz, J. D. & White, D. N. J. (1973). Acta Cryst. A**29**, 93–94.

Goldstein, H. (1950). Classical Mechanics. Cambridge: Addison–Wesley.

Grosse-Kunstleve, R. W. & Adams, P. D. (2002). J. Appl. Cryst. **35**, 477–480.

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

Harris, G. W., Pickersgill, R. W., Howlin, B. & Moss, D. S. (1992). Acta Cryst. B**48**, 67–75.

Howlin, B., Butler, S. A., Moss, D. S., Harris, G. W. & Driessen, H. P. C. (1993). J. Appl. Cryst. **26**, 622–624.

Howlin, B., Moss, D. S. & Harris, G. W. (1989). Acta Cryst. A**45**, 851–861.

Jackson, C. J., Carr, P. D., Liu, J.-W., Watt, S. J., Beck, J. L. & Ollis, D. L. (2007). J. Mol. Biol. **367**, 1047–1062.

Johnson, C. K. (1970). Crystallographic Computing, edited by F. R. Ahmed, pp. 220–226. Copenhagen: Munksgaard.

Johnson, C. K. (1980). Computing in Crystallography, edited by R. Diamond, S. Ramaseshan & K. Venkatesan, pp. 14.01–14.19. Bangalore: Indian Academy of Sciences.

Kuriyan, J. & Weis, W. I. (1991). Proc. Natl Acad. Sci. USA, **88**, 2773–2777.

McCammon, J. A., Gelin, B. R. & Karplus, M. (1977). Nature (London), **267**, 585–590.

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

Painter, J. & Merritt, E. A. (2005). Acta Cryst. D**61**, 465–471.

Painter, J. & Merritt, E. A. (2006a). Acta Cryst. D**62**, 439–450.

Painter, J. & Merritt, E. A. (2006b). J. Appl. Cryst. **39**, 109–111.

Papiz, M. Z., Prince, S. M., Howard, T., Cogdell, R. J. & Isaacs, N. W. (2003). J. Mol. Biol. **326**, 1523–1538.

Prince, E. & Finger, L. W. (1973). Acta Cryst. B**29**, 179–183.

Raaijmakers, H., Törö, I., Birkenbihl, R., Kemper, B. & Suck, D. (2001). J. Mol. Biol. **308**, 311–323.

Rohde, L. A., Ahring, P. K., Jensen, M. L., Nielsen, E., Peters, D., Helgstrand, C., Krintel, C., Harpsøe, K., Gajhede, M., Kastrup, J. S. & Balle, T. (2012). J. Biol. Chem. **287**, 4248–4259.

Šali, A., Veerapandian, B., Cooper, J. B., Moss, D. S., Hofmann, T. & Blundell, T. L. (1992). Proteins, **12**, 158–170.

Scheringer, C. (1973). Acta Cryst. A**29**, 554–570.

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

Schomaker, V. & Trueblood, K. N. (1998). Acta Cryst. B**54**, 507–514.

Sheriff, S. & Hendrickson, W. A. (1987). Acta Cryst. A**43**, 118–121.

Simonetti, A., Marzi, S., Fabbretti, A., Hazemann, I., Jenner, L., Urzhumtsev, A., Gualerzi, C. O. & Klaholz, B. P. (2013). Acta Cryst. D**69**, 925–933.

Sutton, R. B., Ernst, J. A. & Brunger, A. T. (1999). J. Cell Biol. **147**, 589–598.

Tickle, I. & Moss, D. S. (1999). Modelling Rigid-body Thermal Motion In Macromolecular Crystal Structure Refinement. http://people.cryst.bbk.ac.uk/~tickle/iucr99/iucrcs99.html.

Urzhumtsev, A., Afonine, P. V. & Adams, P. D. (2013). Crystallogr. Rev. **19**, 230–270.

Van Benschoten, A. H., Afonine, P. V., Terwilliger, T. C., Wall, M. E., Jackson, C. J., Sauter, N. K., Adams, P. D., Urzhumtsev, A. & Fraser, J. S. (2015). Acta Cryst. D**71**, 1657–1667.

Wilson, M. A. & Brunger, A. T. (2000). J. Mol. Biol. **301**, 1237–1256.

Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). Acta Cryst. D**57**, 122–133.

Winn, M. D., Murshudov, G. N. & Papiz, M. Z. (2003). Methods Enzymol. **374**, 300–321.

Yousef, M. S., Fabiola, F., Gattis, J. L., Somasundaram, T. & Chapman, M. S. (2002). Acta Cryst. D**58**, 2009–2017.

Zucker, F., Champ, P. C. & Merritt, E. A. (2010). Acta Cryst. D**66**, 889–900.