Acta Crystallographica Section B
Structural Science
ISSN 0108-7681

Anisotropic displacement parameters for H atoms using an ONIOM approach

Andrew E. Whitten[a]‡ and Mark A. Spackman[b]*

[a]Department of Chemistry, University of New England, Armidale, NSW 2351, Australia, and [b]Department of Chemistry – M313, University of Western Australia, Crawley, WA 6009, Australia

‡ Present address: Bragg Institute, Australian Nuclear Science and Technology Organization, PMB 1, Menai, NSW 2234, Australia.

Correspondence e-mail: mas@cyllene.uwa.edu.au

Received 27 February 2006
Accepted 1 June 2006

X-ray diffraction data cannot provide anisotropic displacement parameters (ADPs) for H atoms, a major outstanding problem in charge-density analysis of molecular crystals. Although neutron diffraction experiments are the preferred source of this information, for a variety of reasons they are possible only for a minority of materials of interest. To date, approximate procedures combine rigid-body analysis of the molecular heavy-atom skeleton, based on ADPs derived from the X-ray data, with estimates of internal motion provided by spectroscopic data, analyses of neutron diffraction data on related compounds, or *ab initio* calculations on isolated molecules. Building on these efforts, an improved methodology is presented, incorporating information on internal vibrational motion from *ab initio* cluster calculations using the ONIOM approach implemented in *GAUSSIAN03*. The method is tested by comparing model H-atom ADPs with reference values, largely from neutron diffraction experiments, for a variety of molecular crystals: benzene, 1-methyluracil, α-glycine, xylitol and 2-methyl-4-nitroaniline. The results are impressive and, as the method is based on widely available software, and is in principle widely applicable, it offers considerable promise in future charge-density studies of molecular crystals.

1. Introduction

Charge-density analysis is a unique area of crystallography, where the complementary nature of neutron and X-ray diffraction experiments is often exploited, and where theoretical calculations can also provide valuable auxiliary information. Although this information has usually been restricted to electronic structure and properties, increasingly, various aspects of the dynamics of molecules can also be deduced. In a recent commentary, Coppens (2005) argues that, because of recent developments in intense X-ray sources, sensitive-area detectors, cryogenic techniques and computing power, charge densities have 'come of age'. However, that article does not mention a major limitation still faced by the charge-density community while studying organic molecular materials: accurate modelling of the positions and motions of H atoms. This is, of course, due to the fact that X-rays interact with the electron density in the crystal, and the lack of core electrons for hydrogen makes their scattering power poor in comparison with other elements.

Where possible, neutron diffraction data can complement X-ray diffraction data to help pinpoint both the thermal motion and position of H atoms. There are complications associated with the measurement of neutron diffraction data, the most significant being the comparatively low flux of neutron sources with respect to X-ray sources, compounded

by the weak interaction of neutrons with matter, and the combination of these generally demands quite large sample sizes for experiments, especially in comparison with those suitable for X-ray experiments. If a crystal can be grown large enough, then the effects of absorption and extinction are amplified; accurate correction for these effects is not always straightforward, and improper treatment can seriously compromise results of the refinement. Other sources of error include substantial incoherent scattering of neutrons from H atoms, and the possibility of differences between the temperature of X-ray and neutron experiments.

When analysing X-ray diffraction data and modelling electron density deformations due to bonding, the motions and positions of heavier atoms can be determined with similar accuracy and greater precision than from neutron diffraction data. However, spectral truncation of the reflection profile due to the effects of $K\alpha_1,\alpha_2$ splitting can systematically affect the ADPs obtained from X-ray diffraction data (Lenstra *et al.*, 2001; Rousseau *et al.*, 2000). With all these possible effects able to compromise the results obtained, it is not uncommon to find significant differences for heavy-atom ADPs between X-ray and neutron diffraction experiments, sometimes by as much as 50% (Coppens *et al.*, 1984). Various schemes have been proposed to correct the ADPs derived from neutron diffraction experiments, to obtain improved agreement with values derived from X-ray diffraction data (Blessing, 1995), and hence to allow incorporation in a charge-density study. These are based on the assumption that systematic effects apply equally to all atoms, hence the correction factors required to match the X-ray and neutron ADPs of the heavy atoms are also applicable to the H atoms.

Although it is widely recognized that 'no reasonable estimate of the charge-density parameters can be obtained without an adequate description of the thermal motion' (Koritsanszky & Coppens, 2001), the majority of recent charge-density studies model the motion of H atoms isotropically, due to the difficulties inherent in routinely measuring and processing neutron diffraction data. This is a severe approximation, as the amplitudes of bond-bending motions are significantly larger than those associated with bond stretching. In charge-density analysis, charges and isotropic thermal parameters will be correlated to a considerable extent and, in addition, any attempt to refine quadrupole deformations of the electron density will necessarily result in multipole coefficients that are a convolution of both electronic and dynamic effects, as the description of quadrupole deformations and anisotropic thermal motion have the same local symmetry.

Some statistics serve to put this limitation in perspective within the charge-density community. A review of recent publications (1999–2005) reporting experimental charge-density studies reveals that 214 datasets were analysed in detail in that period and, of those, 154 involved organic or molecular systems including H atoms. 79% of the studies involving H atoms treated their thermal motion as isotropic, usually constrained in some way to the motion of the bonded atom, 17% incorporated anisotropic displacement parameters from neutron diffraction experiments on the same compound, and the remainder (<5%) of studies estimated ADPs for H atoms from a combination of rigid-body analysis, and spectroscopic and/or theoretical information. Some recent charge-density studies do not even mention the treatment of thermal motion for the H atoms! This is a remarkable situation, and increasingly unacceptable, given that H atoms play an absolutely vital role in determining electric properties, as well as the crystal packing, of organic molecules. Furthermore, considering the speed and accuracy with which laboratory X-ray and synchrotron data can presently be measured (Luger *et al.*, 2005), the rate-determining step in the present charge-density analysis is no longer experimental: it is the detailed analysis of the experimental measurements, aimed at extracting as much accurate and reliable information as possible from the measurements. In these analyses, much of the recent emphasis has been on topological features of intermolecular interactions, including sophisticated analysis of hydrogen-bond critical-point properties such as the Laplacian, and various energy densities estimated from the experimental $\rho(\mathbf{r})$ (see Munshi & Guru Row, 2005, and references therein), and it has already been demonstrated that use of an isotropic model for H-atom thermal motion has a large effect on all topological features, including those for bonds *not* involving H atoms (Madsen *et al.*, 2004).

This paper outlines our recent work towards a routine approach to the calculation of accurate H-atom ADPs in the absence of neutron diffraction data, using an improved *ab initio* technique applied to molecular clusters in order to take into account the crystal field effects, while keeping computation time to a minimum. We demonstrate that the method is capable of providing remarkably accurate ADPs for H atoms using widely available *ab initio* and crystallographic software, is quite general and is readily applied to crystals containing quite large molecules. The following section summarizes some of the basic concepts and terms in the derivation of ADPs from normal-mode analysis for polyatomic molecules, and this is followed by a brief summary of current approximate methods. Subsequent sections describe the proposed TLS + ONIOM method and its application to X-ray diffraction data for benzene, 1-methyluracil, α-glycine, xylitol and 2-methyl-4-nitroaniline (MNA), which are simple organic molecules spanning a wide range of intermolecular interactions in the crystalline state. We conclude with summary comments, and some suggestions for future improvements.

2. Anisotropic displacement parameters from normal-mode analysis

Much of the pioneering work relating atomic mean-square displacement amplitudes to normal modes of vibration was performed by Cyvin (1968). Because some of that treatment is not always transparent, we summarize the key concepts and expressions in this section. Normal-mode analysis of molecules is treated classically, and the mass-weighted force

constant matrix, **K**, is constructed from second derivatives of the potential energy V,

$$K_{ij} = \frac{\partial^2 V}{\partial q_i \partial q_j}, \tag{1}$$

where $q_1, q_2, q_3, \ldots, q_{3N-1}, q_{3N}$ represent the $3N$ mass-weighted Cartesian coordinates $\sqrt{m_1}x_1$, $\sqrt{m_1}y_1$, $\sqrt{m_1}z_1 \ldots$, $\sqrt{m_N}y_N$, $\sqrt{m_N}z_N$. The force constant matrix can be diagonalized,

$$\mathbf{lKl}^T = \boldsymbol{\lambda}, \tag{2}$$

and the linear transformation **l** represents the normal modes of vibration, transforming between mass-weighted and normal coordinates, **Q** = **lq**. The normal-mode frequencies are obtained from the eigenvalues

$$\omega_k = 2\pi c \nu_k = \lambda_k^{1/2}. \tag{3}$$

While normal-mode analysis of polyatomic molecules is treated classically, the calculation of atomic mean-square displacement amplitudes (MSDAs) is a quantum mechanical problem. The solution of the Schrödinger equation for a harmonic oscillator in terms of the normal coordinate Q_k yields wavefunctions and corresponding energies

$$E_n = \hbar\lambda_k^{1/2}(n + 1/2) = \hbar\omega_k(n + 1/2). \tag{4}$$

From the wavefunctions, the expectation value of the square of the normal coordinate is given by the average over a Maxwell–Boltzmann distribution

$$\langle Q_k^2 \rangle = \frac{\hbar}{2\omega_k} \coth\left(\frac{\hbar\omega_k}{2k_\mathrm{B}T}\right). \tag{5}$$

For a diatomic molecule, the MSDA of the stretching normal mode in Cartesian coordinates is

$$\langle x^2 \rangle = \frac{\hbar}{2\omega\mu} \coth\left(\frac{\hbar\omega}{2k_\mathrm{B}T}\right), \tag{6}$$

where μ is the effective mass. For a polyatomic molecule, the equivalent expression for the kth normal mode is

$$\delta_{kk} = \langle Q_k Q_k \rangle = \frac{\hbar}{2\omega_k} \coth\left(\frac{\hbar\omega_k}{2k_\mathrm{B}T}\right), \tag{7}$$

and for a collection of normal modes this can be summarized in matrix fashion

$$\boldsymbol{\delta} = \langle \mathbf{QQ}^T \rangle = \mathrm{diag}(\delta_1, \delta_2, \ldots, \delta_n). \tag{8}$$

$\boldsymbol{\delta}$ is necessarily diagonal as **Q** refers to normal coordinates. To express the molecular mean-square Cartesian amplitude matrix $\langle \mathbf{xx}^T \rangle$ in terms of the normal coordinates, (9) is used to relate the normal coordinates to the Cartesian coordinates

$$\mathbf{Q} = \mathbf{lq} = \mathbf{lm}^{1/2}\mathbf{x}, \tag{9}$$

from which it follows that

$$\langle \mathbf{QQ}^T \rangle = \mathbf{lm}^{1/2} \langle \mathbf{xx}^T \rangle \mathbf{m}^{1/2}\, \mathbf{l}^T, \tag{10}$$

and, because **l** is a unitary transformation ($\mathbf{l}^T = \mathbf{l}^{-1}$), then upon rearrangement of (10),

$$\begin{aligned} \langle \mathbf{xx}^T \rangle &= \mathbf{m}^{-1/2}\, \mathbf{l}^{-1} \langle \mathbf{QQ}^T \rangle (\mathbf{l}^{-1})^{-T} \mathbf{m}^{-1/2} \\ &= \mathbf{m}^{-1/2}\, \mathbf{l}^T \boldsymbol{\delta} \mathbf{l} \mathbf{m}^{-1/2}. \end{aligned} \tag{11}$$

This matrix contains both the atomic and interatomic MSDAs. The ADPs form the 3×3 diagonal blocks of this matrix, and the off-diagonal 3×3 blocks are the interatomic or correlation mean-square amplitudes, and are not determinable from a diffraction experiment (Bürgi & Capelli, 2000).

There are other equivalent variations of this, and one of relevance to the results that follow is that used by the *GAUSSIAN03* package (Frisch *et al.*, 2004; Ochterski, 1999). Frequencies are obtained by the procedure described above, but, instead of reporting normalized mass-weighted Cartesian displacements **l**, normalized Cartesian displacements **e** are obtained by the transformation

$$\mathbf{e} = \mathbf{N}\mathbf{m}^{-1/2}\, \mathbf{l}^T. \tag{12}$$

N is a diagonal matrix of normalizing factors, each element related to the effective mass of the given mode

$$\mu_k = N_k^2. \tag{13}$$

It then follows that the diatomic case, (6), can be generalized to apply to polyatomic molecules

$$\langle \mathbf{xx}^T \rangle = \mathbf{e}\boldsymbol{\mu}^{-1}\boldsymbol{\delta}\mathbf{e}^T. \tag{14}$$

3. Approximate methods combining estimates of internal and external motions

It was recognized by Hirshfeld (1976) that information from both X-ray diffraction and IR spectroscopy could be combined to estimate ADPs for hydrogen nuclei, and this is the basis of most methods that have been used for this purpose. The assumption made in approaches of this kind is that the external and internal motions are well separated into low- and high-frequency vibrational modes of a molecule in a crystal, so that the internal displacement can be expressed as a sum of the internal high frequency, and the external displacements as a sum of the low-frequency modes of vibration

$$U_{ij} = U_{ij}^{\mathrm{internal}} + U_{ij}^{\mathrm{external}}. \tag{15}$$

The most common method for estimating the external contribution to the total motion of a molecule is the 'rigid body' or TLS analysis (Cruickshank, 1956; Dunitz, Schomaker & Trueblood, 1988; He & Craven, 1993; Schomaker & Trueblood, 1968, 1998). This method is based on the assumption that the crystal field has hindered the rotations and translations of a free molecule such that the molecule still moves as a rigid body, but the motion is now composed of harmonic translations and librations. Hence, a rigid body with translational and librational degrees of freedom can be fitted to the observed ADPs of any subset of atoms in the molecule, to yield a model for the external motion of a molecule in a crystal, which can subsequently be used to estimate the external motion for all atoms in the structure. Based on a detailed breakdown of 15 K neutron ADPs for C_6D_6 (Dunitz,

Maverick & Trueblood, 1988), this approach is expected to work well for molecules such as benzene, and we show below that this is indeed the case. However, we also recognize that for most molecules of interest in charge-density analysis the low-frequency vibrations include contributions from both internal and external motions of a molecule, and the more complex segmented rigid-body analysis, for example the use of attached rigid groups (Schomaker & Trueblood, 1998), will need to be invoked in many instances. Nevertheless, we focus in the present work on the relatively simple TLS analysis, with the aim of establishing a benchmark for future investigations.

Important examples of this kind of analysis involve choosing frequencies from IR and Raman spectroscopic data that correspond to assumed internal modes of vibration (Destro *et al.*, 1989, 2000; May *et al.*, 2001; Roversi *et al.*, 1996; Roversi & Destro, 2004). For example, spectroscopic frequencies are assigned to stretching and in- and out-of-plane bending modes for specific hydrogen nuclei. The direction of the vector **e** in (14) is assumed, and since the mass of the hydrogen (or deuterium) nucleus is known, as is the frequency for each mode, the mean-square displacements in three orthogonal directions can be determined for each hydrogen nucleus, and ADPs are then calculated from a sum of the external (estimated from a TLS model) and internal contributions. There are, however, two problems with this method. Firstly, the motion of an atom in any one direction is associated with a single vibrational mode, whereas in reality a combination of many modes make up the motion of an atom in any one direction. Equation (14) illustrates that in effect, an eigenvector **e** may distribute its corresponding mean-square amplitude δ between several contributing atoms and also that more than one normal mode may contribute to the displacements of a given atom in a given direction. For this reason the amplitude of vibration will in general be biased, *i.e.* over- or underestimated, depending on the frequency chosen to be most representative of atomic motion in a given direction. A second problem with this type of analysis is that the internal contributions to the heavy-atom ADPs are not normally subtracted before the TLS analysis, which means that the rigid-body model obtained is a mixture of both internal and external motions. Even though these internal contributions are small for heavy atoms, the components of the libration and translation tensors may be significantly affected, and, although the fitted ADPs for the heavy atoms may differ by a small amount, the effect on the H atoms lying on the periphery of the molecule may be quite large. The effect of the internal-motion contribution to rigid-body fits for C_6D_6 has been analysed by Capelli *et al.* (2000), who concluded that the calculated translation and libration tensors are strongly dependent on the internal motion model.

To a good approximation the high-frequency normal modes in molecular crystals describe internal molecular motions and are similar to those for gas-phase molecules; this greatly simplifies the task of calculating internal contributions to the ADPs of atoms in a molecular crystal, as extraction of both normal modes and frequencies from *ab initio* calculations is a relatively routine procedure, as outlined above. Estimates of the internal motion derived from calculations on isolated molecules can therefore be directly transferred to molecules in a crystal, and the capacity for this analysis is now implemented in the *XD* suite of programs (Koritsanszky *et al.*, 2003). Several examples of this approach to the estimation of ADPs for H atoms have been published, with varying levels of success (Flaig *et al.*, 1998; Madsen *et al.*, 2003, 2004; Oddershede & Larsen, 2004). This approximation, however, works well only for small aromatic or rigid systems that are not involved in strong intermolecular bonding, which almost precludes its use with many interesting compounds. Problems are encountered where the conformation of the free molecule differs significantly from that in the crystal, in which case the similarities between the gas phase and the crystal are lost. The advantage of the *ab initio* approach over all spectroscopic approaches is that the displacement vectors of the normal modes are known and need not be assumed.

One further method that has been used to estimate mean-square displacements for internal modes utilizes neutron diffraction data (Madsen *et al.*, 2003, 2004; Weber *et al.*, 1991). This procedure takes the internal contribution of the ADPs for hydrogen nuclei to be the difference between those determined from a rigid-body model fitted to the experimental heavy-atom ADPs, and the experimental ADPs for hydrogen. The principal components are then assigned to bond stretching or bending modes, and grouped according to the environment of the H atom. For each given environment the internal amplitudes of vibration can be averaged and applied to other molecules. This method has been shown to be quite successful; however, the results are affected by the fact that internal motion of the heavy atoms has not been taken into account in the rigid-body analysis.

Bürgi and co-workers have recently described an elegant method for studying the dynamics of molecules in crystals through the analysis of multi-temperature neutron diffraction data (Bürgi & Förtsch, 1999; Bürgi & Capelli, 1999, 2000; Bürgi, 2000; Bürgi *et al.*, 2001; Capelli *et al.*, 2000). The normal modes derived from this approach incorporate both internal and external degrees of freedom, and can also be used to estimate the ADPs for any atom at any intermediate temperature (Bürgi *et al.*, 2002). The procedure is a modification of the Einstein approximation, where it is considered that molecules, not only atoms, vibrate in a mean field of their neighbours (Bürgi & Capelli, 2000). Within this approximation, the temperature dependence of the ADPs is well understood, the only difference between the ADPs of a molecule in a crystal at two temperatures being the contribution from low-frequency modes. The high-frequency modes are essentially invariant to changes in the temperature, while the low-frequency modes are temperature independent at very low temperatures, and display a linear dependence on temperature in the high-temperature limit. There is insufficient information in conventional X-ray or neutron diffraction experiments at a single temperature to separate the internal and external contributions to the ADPs, but data from two widely separated temperatures are sensitive to this information. In analyses of this kind the eigenvectors and frequencies

for the low-frequency modes are refined, along with an overall internal contribution to the motion for each atom. In practice, there are still too many parameters to determine, and often many constraints must be introduced (for example, internal amplitudes of vibration are the same for similar atom types), and the motion is analysed in terms of a segmented rigid body. The procedure also overcomes the indeterminacy issues associated with TLS analysis, as it allows ADPs at different temperatures to be expressed using the same displacement vectors, frequencies and internal contributions to the motion. Details can be found elsewhere (Bürgi & Capelli, 2000), but the analysis is essentially the same as a segmented rigid-body analysis except that the temperature dependence of the low-frequency modes and the contribution to the motion of the internal modes of vibration are explicitly included.

4. The TLS + ONIOM approximation

A common limitation of most methods discussed in the previous section is that they cannot be applied in a routine fashion to X-ray diffraction data measured at a single temperature. For this reason, modern *ab initio* techniques offer an attractive method for calculation of internal modes and frequencies. Previous theoretical calculations (Luo *et al.*, 1996; Madsen *et al.*, 2003) have suggested that the environment of larger non-rigid molecules is responsible for markedly altering the magnitude of low-frequency internal modes. Gas-phase molecules undergo low-frequency large-amplitude torsions, while in the presence of the crystal field these motions will typically be hindered, in effect raising the frequency of vibration, which in turn means the amplitude of vibration is smaller. There is no doubt that crystal field effects can be modelled by clusters of molecules or atomic fragments surrounding a central molecule, but computing time rapidly increases with the size of the system, and most interesting molecules studied by charge-density analysis are moderately sized (>15 atoms), hence full geometry optimizations of even small clusters of molecules are prohibitive even at modest levels of theory.

Alternative methods allow different regions of a system to be modelled at different levels of theory, keeping computation time to a minimum for large structures (Morokuma, 2003). So-called mixed quantum-mechanics/molecular-mechanics (QM/MM) implementations such as 'integrated molecular orbital molecular mechanics' [IMOMM (Maseras & Morokuma, 1995)] and the more versatile 'our own *N*-layer integrated molecular orbital molecular mechanics' [ONIOM (Dapprich *et al.*, 1999; Svensson *et al.*, 1996; Vreven *et al.*, 2003] calculations are used where information is sought on a specific region of a large system, and the surroundings have a small but important effect on the geometry, or other aspects of the system. The ONIOM method is more versatile than conventional QM/MM methods, as any combination of calculation types such as density functional, Hartree–Fock and molecular mechanics can be used in the same calculation. This type of analysis has not yet been used to investigate the motion of molecules in a crystal, but appears ideal for modelling the effect of the crystal environment on a central molecule. As in previous work applied to single molecules, an *ab initio* method can be used to model a single molecule, while the influence of the surroundings is modelled with a lower level of theory, even molecular mechanics.

To explore the possibilities offered by methods of this kind, two-layer ONIOM calculations were performed on molecular clusters for several molecules which have been the subject of charge-density analyses. These calculations are in the same vein as recent work on amino acids by Zheng *et al.* (2004), but with a number of significant differences. Initial molecular geometries were taken from experimental neutron or charge-density studies, and all surrounding molecules with an atom closer than 8 Å to any atom in the central molecule were included in the cluster. Molecules in the outer layer were modelled using the molecular mechanics universal force field (UFF) (Rappé *et al.*, 1992) and the central molecule was described at the Hartree–Fock level of theory. The basis set chosen for the central molecule is based on a DZP set due to Thakkar *et al.* (1993), supplemented with additional *d*-type polarization functions (exponents 0.75, 0.80, 0.85 and 0.90 a.u., for C, N, O and F, respectively), a *p*-type polarization function for H (exponent 1.0 a.u.), and diffuse *s*- and *d*-type functions on C, N, O and F, and *s*- and *p*-type functions on H (Dougherty & Spackman, 1994). Atomic charges for the outer layer were included in the quantum mechanical Hamiltonian, so-called electronic embedding (Vreven *et al.*, 2006), while the molecular mechanics layer provided the appropriate repulsion potentials to replicate the environment in the crystal. The atomic charges were determined from a least-squares fit to the electric field around the molecule generated from periodic Hartree–Fock calculations as outlined in detail elsewhere (Whitten, McKinnon *et al.*, 2006). Previous studies have examined the influence of including point charges to model the crystal field in urea (Rousseau *et al.*, 1998, 1999), where charges on surrounding molecules were altered in a self-consistent manner, such that they matched the Mulliken charges of the central molecule. In the present work we incorporate atomic charges that closely reproduce the electric field from a crystal calculation at points around the central molecule in the cluster, and these charges are not altered during the calculation.

The geometry of the central molecule (high-level layer) was optimized, while the geometries of all surrounding molecules were kept fixed. In general, optimized ONIOM molecular geometries and conformations were in excellent agreement with experimental crystal geometries in all cases, although systematic differences in bond lengths were obtained that are consistent with corresponding differences for isolated molecules at this level of theory.[1] For example, ONIOM bond lengths for X—H bonds are typically shorter by 0.01–0.02 Å than experimental values, single O—C bonds are shorter by ~0.01 Å, but single C—C and C—N bonds are longer by

[1] See, for example, comparisons between Hartree–Fock/6-31G** results and experiment for bond lengths and bond angles in the NIST Computational Chemistry Comparison and Benchmark Database: http://srdata.nist.gov/cccbdb/.

about the same amount, and C—C and C—N bonds in conjugated ring systems are shorter by ~0.01 Å. Bond angles are reproduced to within a few degrees, even for C—O—H bonds, while agreement for torsion angles is slightly worse, and in xylitol the difference is as much as 14° for C2—C1—O1—H11. The ONIOM model also predicts the NH_2 group in MNA to be significantly more pyramidal than observed in the crystal (although the near-planar crystal geometry could represent the average over an inversion motion about N1).

Normal-mode analysis was then performed on the geometry-optimized molecule, requiring the normal modes and frequencies of the central molecule to be separated from those in the rest of the cluster. This was achieved by specifying isotopic masses of 10^5 a.m.u. for all atoms in the outer layer (low-level layer). This means that the resulting frequencies were obtained due to a mean (static) field of neighbours and yielded in all cases $3N$ modes of vibration. Preliminary analysis indicated that it was not always possible to unambiguously classify normal modes as internal or external; with the exception of benzene, many low-frequency modes are evidently a mixture of the two, which means that the ONIOM calculation provides an approximation to the intramolecular mean-square amplitudes, the first term in (15). While the high-frequency modes appear to yield a realistic contribution to the ADPs, the force constants describing the interaction between the central and surrounding molecules appeared to be overestimated, and their contribution to the ADPs underestimated, and two options were explored to compensate for this limitation.

The first was to disregard the six lowest-frequency (external) modes and scale the frequencies of the remaining $3N - 6$ modes by a common factor of 0.90 (Scott & Radom, 1996) to estimate the internal contribution to the ADPs of the heavy atoms. These contributions were then subtracted from the experimental X-ray ADPs of the heavy atoms, and a TLS fit performed on the results. ONIOM estimates of the internal contributions to the ADPs for H atoms were then added to those implied by this TLS model to yield what we refer to as TLS + ONIOM ADPs for H atoms. ONIOM-derived internal contributions to the ADPs are without doubt superior to those provided by an isolated molecule calculation, but for flexible molecules they may lack an unknown, usually small, contribution from internal degrees of freedom but contain another small but unknown contribution from external degrees of freedom.

Another attractive option relies entirely on ONIOM normal-mode frequencies, but scales internal- and external-mode frequencies by different amounts. The practice of scaling normal-mode frequencies obtained from *ab initio* calculations is now common (Scott & Radom, 1996), and it attempts to correct for basis set inadequacies, lack of electron correlation and anharmonicity effects. For normal *ab initio* calculations this is equivalent to scaling the entire Hessian matrix by a single scale factor, but a Hessian derived from an ONIOM calculation includes contributions from different levels of theory (in the present case both Hartree–Fock and molecular mechanics), and it is not obvious how to scale the Hessian (or normal-mode frequencies) to account for this difference. It is clear, however, that internal modes (which depend largely on the *ab initio* portion of the calculation) and external modes (which depend largely on the molecular mechanics description) deserve to be scaled differently.

The suitability of scaling of the low-frequency modes can be tested for deuterated benzene (C_6D_6) as multi-temperature neutron diffraction data on this compound have been analysed in terms of a modified Einstein model, yielding the frequencies of the external modes (Capelli *et al.*, 2000). Fig. 1 compares the external-mode frequencies obtained from this analysis of multi-temperature neutron diffraction data with the frequencies of essentially the same six modes obtained from the ONIOM cluster calculation for C_6D_6. It can be seen from the figure that theoretical frequencies of external modes are systematically much greater than those obtained from detailed analysis of multi-temperature neutron diffraction data, suggesting that, in addition to scaling the high-frequency modes by 0.90, the external modes for benzene should be scaled by ~0.55 in order to yield the best agreement with the experimental ADPs. The ADPs resulting from this scaling, for both C and D atoms, are compared with ADPs derived from neutron diffraction data in Fig. 2, and the agreement is seen to be excellent; the root-mean-square deviation of the ADPs for all atoms is 4.8% at 15 K and 7.5% at 123 K. Benzene clearly represents an almost ideal case for this approach, largely because there are no low-frequency internal molecular modes of vibration, and hence low- and high-frequency modes from the ONIOM calculation are well separated (by 318 cm^{-1} for C_6D_6 and 364 cm^{-1} for C_6H_6). For other compounds where the lowest-frequency modes almost invariably mix internal and external motions, it was found that scaling all low-frequency modes by a constant yielded comparatively poor agreement of ADPs with reference results, and for this reason the TLS + ONIOM approach was the only one pursued further for molecules more complicated than benzene.

5. Applications of the TLS + ONIOM model to molecular crystals

We present results for five different molecular crystals, each of which has been the subject of a detailed charge-density

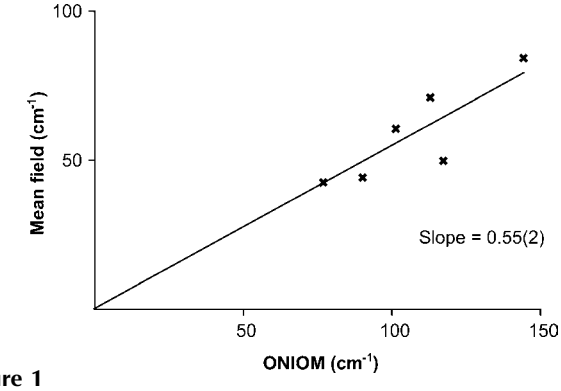

Figure 1
Comparison of external-mode frequencies of C_6D_6 obtained from analysis of multi-temperature neutron diffraction data (Capelli *et al.*, 2000) with theoretical values from an ONIOM cluster calculation.

analysis. In each of these charge-density studies, H-atom ADPs were obtained from either neutron diffraction data or some approximate method, and those values will be used as a reference for comparison with the present TLS + ONIOM model ADPs. The internal contributions to the model ADPs have been calculated from the $3N - 6$ highest-frequency modes obtained from the ONIOM cluster calculation with *GAUSSIAN03* (Frisch *et al.*, 2004), with frequencies scaled by a common factor of 0.90. These were then subtracted from the heavy-atom ADPs obtained from the X-ray diffraction data, and a TLS model fitted to the corrected ADPs using the *THMA*11 program (Schomaker & Trueblood, 1998). For H atoms, the ONIOM internal contributions to the ADPs were then added to those obtained from the TLS fit, yielding approximate TLS + ONIOM ADPs; for this purpose we used either our own in-house software or the *XDVIB* module of the *XD* package (Koritsanszky *et al.*, 2003). Benzene is an exception to this procedure, as it is not possible to uniquely fit a rigid body to the heavy-atom skeleton, for which the atoms lie close to a conic section. Instead, the $3N - 6$ highest-frequency modes were again scaled by a common factor of 0.90, but the remaining six modes were scaled by the factor 0.55 (previously determined to be optimum for C_6D_6), from which ADPs were then calculated.

Where neutron data were available at approximately the same temperature as the X-ray experiment, reference H-atom ADPs were obtained by adjusting the neutron values to best fit the X-ray ADPs for heavy atoms, $\mathbf{U}_X = q\mathbf{U}_N + \Delta\mathbf{U}$, using Blessing's approach as coded in *UIJXN* (Blessing, 1995). Quantitative comparisons between calculated and reference ADPs have previously been made using a conventional least-squares statistic based on squares of differences, as in the program *THMA*11 (Schomaker & Trueblood, 1998), or based on absolute values of differences (Madsen *et al.*, 2004).

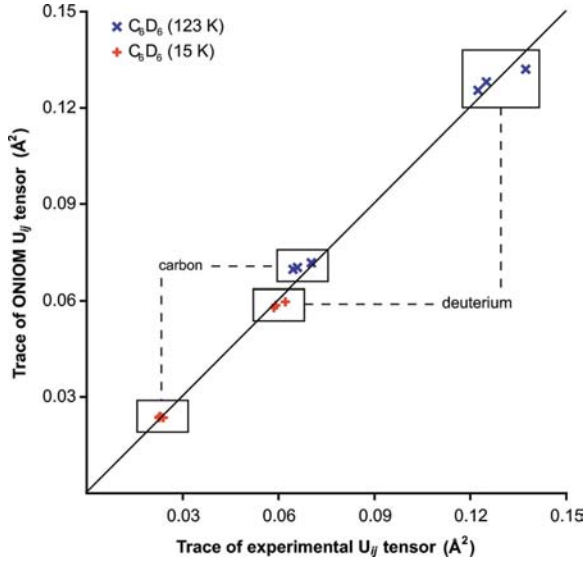

Figure 2
Comparison between ADPs obtained from scaled ONIOM frequencies and those from neutron diffraction for C and D atoms in C_6D_6 at 15 and 123 K (Jeffrey *et al.*, 1987). The line on the graph has unit slope.

However, agreement statistics such as these, based directly on the unique values of U_{ij}, are not necessarily independent of the choice of cell axes. A more appropriate measure of agreement can be defined in terms of the overlap between the two probability density functions (pdfs) in direct space. The ADP tensor $\mathbf{U}$ represents a normalized pdf in real space of the form (Willis & Pryor, 1975)

$$p(\mathbf{u}) = \left(\frac{\det \mathbf{U}^{-1}}{8\pi^3}\right) \exp\left(-\frac{1}{2}\mathbf{u}^T\mathbf{U}^{-1}\mathbf{u}\right), \qquad (16)$$

where $\mathbf{u}$ generally refers to non-orthogonal cell axes. To evaluate an appropriate normalized measure of overlap it is convenient to transform $\mathbf{U}$ to a Cartesian system, and for two different ADP tensors $\mathbf{U}_1$ and $\mathbf{U}_2$ expressed with respect to these (arbitrary) Cartesian axes the integral is given by

$$R_{12} = \int \left[p_1(\mathbf{x})p_2(\mathbf{x})\right]^{1/2} \mathrm{d}^3\mathbf{x} = \frac{2^{3/2}\left(\det \mathbf{U}_1^{-1}\mathbf{U}_2^{-1}\right)^{1/4}}{\left[\det\left(\mathbf{U}_1^{-1} + \mathbf{U}_2^{-1}\right)\right]^{1/2}}. \qquad (17)$$

Because the pdfs are normalized, this overlap integral has the desired property that if $\mathbf{U}_1 = \mathbf{U}_2$, $R_{12} = 1.0$. In practice, values of R_{12} are only slightly less than 1.0, and a more discriminating similarity index was introduced, $S_{12} = 100(1 - R_{12})$, which conveniently describes a percentage difference between the two pdfs represented by $\mathbf{U}_1$ and $\mathbf{U}_2$. This index was computed for each H atom to assess agreement between reference ADPs and our present TLS + ONIOM model results; for molecules, we also report values averaged over the H atoms in the molecule, $\bar{S}_{12}$.

Before discussing results for each molecular crystal in detail, we examine the validity of a key assumption in our present approach, namely the question of rigidity *versus* non-rigidity for these molecules, based on the analysis of heavy-atom ADPs. This is most easily answered by examination of the matrix of MSDA differences between all pairs of atoms along interatomic directions

$$\Delta_{A,B} = \mathbf{n}^T\mathbf{U}_A\mathbf{n} - \mathbf{n}^T\mathbf{U}_B\mathbf{n}, \qquad (18)$$

where $\mathbf{n}$ is the unit vector along the A–B direction. As discussed elsewhere (Dunitz, Maverick & Trueblood, 1988; Dunitz, Schomaker & Trueblood, 1988; Rosenfield *et al.*, 1978), a necessary (but not sufficient) condition for a rigid body requires elements of this matrix to be close to zero for all intramolecular atom pairs (*i.e.* including non-bonded pairs). According to Hirshfeld (1976), matrix elements for bonded atoms should be less than 10 pm^2 for typical bonds in organic molecules, and we use this criterion to identify possible non-rigidity. For all molecules subjected to a TLS analysis, the r.m.s. value of $\Delta_{A,B}$ is *smaller* after subtracting internal contributions from the X-ray ADPs, reinforcing the importance of correcting X-ray ADPs for internal motions before performing a rigid-body analysis;[2] following this correction, $\Delta_{A,B}$ matrices for 1-methyluracil, glycine and MNA have no elements greater than 9 pm^2. For xylitol, the largest matrix

[2] The differences are often substantial: 1-methyluracil, from 4 to 2 pm^2; glycine, from 3 to 2 pm^2; xylitol, from 11 to 9 pm^2; and for MNA, from 18 to 4 pm^2.

elements are observed for atom pairs involving O4 and O5 (O4–C7: −14 pm^2, O4–O5: −25 pm^2, O3–O5: −20 pm^2 and O3–O4: 18 pm^2), suggesting that there is significant internal motion about the C—C bonds in that molecule, and the results presented below for this molecule should be assessed with this in mind.

5.1. 1-Methyluracil

Single-crystal neutron diffraction data have been measured for 1-methyluracil at 15, 60 and 123 K by McMullan & Craven (1989), and the refined ADPs were analysed in terms of a rigid-body model. The data, although of good quality, required an anisotropic extinction correction, with the worst attenuation of intensity in the 15 K data set being almost 65%. Systematic differences were found between the observed and TLS model values of U_{33} for each nucleus, and this was attributed to internal ring puckering vibrations. However, a subsequent charge-density study (Klooster *et al.*, 1992) also revealed substantial differences between the X-ray refined values of U_{33} for heavy atoms, and those from McMullan & Craven. Quite recently, hydrogen ADPs from the 15 K neutron experiment were compared with those obtained from rigid-body analysis of X-ray data measured at 21 K (Roversi & Destro, 2004), with estimates of internal motion for H atoms obtained from solid-state spectroscopy (Lewis *et al.*, 1984; Szczesniak *et al.*, 1985). The general agreement was very good, with major discrepancies being U_{33} for H3 and U_{11} for H12 (see Fig. 3 for atom numbers). This latter analysis assumed that the motion due to the internal vibrations of the heavy nuclei is negligible (*i.e.* those contributions were not subtracted before performing the TLS fit), and only selected internal vibrational modes were included in the calculation.

Based on a comparison between the frequencies obtained from an ONIOM cluster calculation and those computed for a free molecule, it could be argued that the effect of the crystal field on 1-methyluracil is small, as most frequencies are close in magnitude, but an important exception is observed for the methyl group, which undergoes large-amplitude low-frequency vibrations (or hindered rotation) in the absence of a crystal environment. A free molecule *ab initio* calculation on 1-methyluracil using the same level of theory and basis set as the high-level layer of the ONIOM calculation estimates the frequency of this mode as approximately 86 cm^{-1}, whereas the influence of the crystal field shifts it to approximately 200 cm^{-1}, and this will clearly have a marked influence on the ADPs as the low-frequency modes make the greatest contributions to the atomic displacements. The notation ε_{11}, ε_{22} and ε_{33} will be used from this point to describe the mean-square displacement amplitudes of H-atom stretching, in- and out-of-plane bending modes of motion, respectively. This assignment is well defined for planar (or near planar) molecules, but it becomes ambiguous for other chemical groups. For a methyl H atom, the plane is defined to be through the C–H bond, bisecting the angle subtended by the other two H atoms and the methyl C atom. The plane for methylene H atoms is defined to be through both H atoms and the bonded C atom. There is no unambiguous definition possible for in- and out-of-plane components for methine H atoms.

Table 1
Experimental IR and Raman vibrational frequencies, ν (cm^{-1}) (Szczesniak *et al.*, 1985), and corresponding mean-square displacement amplitudes, ε (Å^2 × 10^4), for three principal vibrational modes for H atoms in 1-methyluracil.

For each atom the first row reports results from Roversi & Destro (2004), and the second row those from the present ONIOM calculation (as described in the text, frequencies have been scaled by 0.90).

	ν_{11}	ν_{22}	ν_{33}	ε_{11}	ε_{22}	ε_{33}
H3	3114	1436	832	58	117	190
	3154	1462†	896	64	125	201
H5	3040	1055	805	60	159	208
	3080	1146†	767†	66	150	240
H6	3040	1228	921	60	136	182
	3030	1266†	1032	66	138	230
H11	2960	1424	150	61	117	353
	2938†	1396†	202†	72	155	392
H12	2960	1424	150	61	117	353
	2938†	1396†	202†	64	171	390

† These values represent averages over many similar modes.

Table 1 shows that the scaled ONIOM vibrational frequencies obtained for 1-methyluracil agree very well with experimental solid-state IR and Raman results and, as expected, almost all the stretching and in- and out-of-plane vibrational amplitudes reported by Roversi & Destro (2004) are smaller than the present ONIOM results, which include contributions from all vibrational modes. For U_{11} on H12, Roversi & Destro (2004) reported a large difference between their computed value and that derived from neutron diffraction data, and this component contains contributions from mainly the in- and out-of-plane vibrations. Table 1 shows that both these components are substantially underestimated by the spectroscopic approach, compared with the present ONIOM result. This discrepancy mainly arises from the low-frequency rocking modes of the methyl groups and asymmetric deformations of the methyl H atoms, which cannot be modelled using the limited spectroscopic information. The ONIOM calculation suggests that there are a variety of motions that contribute to the motion of the methyl hydrogen

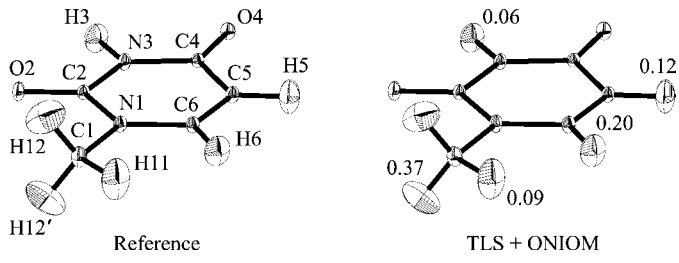

Figure 3
ORTEP representation and atom labelling for 1-methyluracil at 21 K (99% probability level); H12 and H12′ are related by a mirror plane. Reference values are calculated from neutron ADPs (McMullan & Craven, 1989) adjusted against X-ray ADPs (Roversi & Destro, 2004) using *UIJXN* (Blessing, 1995). For the TLS + ONIOM model, S_{12} indices are given for each H atom; $\bar{S}_{12}$ = 0.17.

nuclei between 160 and 420 cm^{-1}, and the complicated nature of these modes cannot be accounted for in the spectroscopic approach as the normal modes are unknown. The other significant discrepancies reported by Roversi & Destro (2004) are for U_{33} on H3 and H6. These differences are clearly not caused by the description of the internal motion, as the contribution to the internal motion is similar for both the ONIOM and spectroscopic methods.[3]

At 21 K the TLS + ONIOM model reproduces the reference ADPs of the H atoms very well, even for those in the methyl group. Fig. 3 shows that the principal directions and sizes of the thermal ellipsoids are similar in all cases. The mean similarity index, $\bar{S}_{12} = 0.17$, is small, and compares favourably with a value of $\bar{S}_{12} = 0.39$ obtained for a comparison between the estimated and reference ADPs reported by Roversi & Destro (2004; Table 2 in that work). Although they are not included in the TLS + ONIOM model, we note that the six lowest-frequency ONIOM normal modes are not realistic, and their use leads to problems (for 1-methyluracil) in the values of U_{33}, which correspond to out-of-plane deformations, causing large variations between reference and the ONIOM method at higher temperatures. The causes of the unrealistic normal modes are probably numerous, but it is likely that the UFF force field used to describe the intermolecular interactions is poorly modelling the actual effects.

5.2. Benzene

Benzene is a fortunate case with respect to this kind of analysis as it is small, rigid, and the internal- and external-mode frequencies are well separated. On the other hand, TLS analysis of the heavy-atom skeleton of benzene does not provide a unique answer, hence it is not possible to use a rigid body to obtain estimates of the external motion. In a recent charge-density study on benzene (Bürgi *et al.*, 2002), information obtained from the analysis of multi-temperature neutron diffraction data on C_6D_6 (Capelli *et al.*, 2000) was used to calculate the ADPs for H atoms in C_6H_6. External motion for C_6H_6 was estimated by reconstruction of the force constant matrix from the normal modes and frequencies obtained from the analysis of multi-temperature neutron diffraction data on C_6D_6, and then application of a new isotopic composition to obtain a new set of normal modes and frequencies at the temperature of the X-ray diffraction experiment (Bürgi *et al.*, 2002). Since the mass dependence of the combination of internal modes is complicated, estimates of the internal motion were taken from analysis of multi-temperature neutron data on the benzene complex $AgClO_4.C_6H_6$ (Bürgi & Capelli, 1999; McMullan *et al.*, 1997).

Internal contributions to the ADPs of carbon, deuterium and hydrogen for C_6H_6 and C_6D_6 are listed in Table 2, obtained from both spectroscopy and neutron diffraction, and these are compared with the results obtained from the present ONIOM calculations. The ONIOM values are in agreement with spectroscopic values, indicating that the surroundings in this case are not significantly influencing the magnitudes of the internal vibrations. Largely because of this, we see from the thermal ellipsoids and the similarity indices in Fig. 4 that simple scaling of the external-mode frequencies by a common scale factor (0.55) yields ADPs in excellent agreement with the reference neutron values. The value of $\bar{S}_{12} = 0.16$ is slightly smaller than that found for 1-methyluracil, largely because of the use of this optimum scale factor.

Table 2
Average intramolecular mean-square displacement amplitudes ($Å^2 \times 10^4$) for carbon and H/D atoms in C_6D_6 and C_6H_6.

The first two rows for each compound are from Bürgi & Capelli (1999) and the last row lists results from ONIOM calculations on a cluster of benzene molecules (as described in the text, frequencies have been scaled by 0.90).

	C			H/D		
	ε_{11}	ε_{22}	ε_{33}	ε_{11}	ε_{22}	ε_{33}
C_6H_6						
Neutron diffraction	14 (1)	11 (1)	11 (1)	68 (2)	124 (2)	171 (3)
Spectroscopic force field	9	12	14	61	130	202
ONIOM	13	8	14	64	132	196
C_6D_6						
Neutron diffraction	14 (1)	7 (1)	15 (1)	52 (1)	83 (1)	110 (2)
Spectroscopic force field	13	8	16	44	89	133
ONIOM	13	9	16	46	90	129

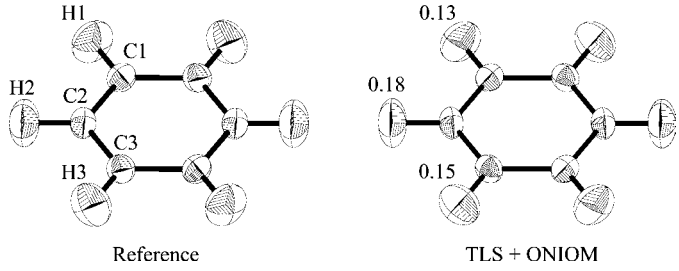

Figure 4
ORTEP representation and atom labelling for C_6H_6 at 110 K (99% probability level); unlabelled atoms are related by inversion. Reference values are calculated from mass-adjusted multi-temperature neutron diffraction data for C_6D_6 (Bürgi *et al.*, 2002). For the TLS + ONIOM model, S_{12} indices are given for each H atom; $\bar{S}_{12} = 0.16$.

5.3. α-Glycine

Whereas benzene presented a specific problem due to indeterminacies related to the rigid-body fitting procedure, problems for glycine involve the estimates of internal motion, specifically from theory. As for the previous study on 1-methyluracil by Roversi & Destro (2004), a previous charge-density study on α-glycine (Destro *et al.*, 2000) used the combination of specific spectroscopic frequencies (Machida *et al.*, 1977) and assumed normal-mode displacement patterns to reconstruct the internal motion for H atoms. Calculation of theoretical frequencies and internal modes derived from *ab initio* calculations on an isolated molecule is not possible

[3] Supplementary data, including complete reference and TLS + ONIOM H-atom ADPs for 1-methyluracil, benzene, glycine, xylitol and MNA, are available from the IUCr electronic archives (Reference: LB5003). Services for accessing these data are described at the back of the journal.

for glycine due to the intermolecular bonding environment in the crystal, where glycine is zwitterionic and forms a complicated hydrogen-bonded structure. Geometry optimization performed on an isolated molecule yields a lowest energy structure that is not zwitterionic. Compounds of this type require the crystalline environment to keep the molecule in the zwitterionic form, and application of the present ONIOM approach provides insight into the general applicability of this method for systems such as this.

Room-temperature neutron diffraction data have been reported for α-glycine (Jonsson & Kvick, 1972), but for obvious reasons the ADPs are not applicable to the 23 K charge-density study of Destro *et al.* (2000), hence the need for determination of approximate ADPs for H atoms. It is apparent that the estimated ADPs at 23 K reported by Destro *et al.* (2000) are a good approximation to the actual ADPs at that temperature as their charge-density analysis yielded quadrupole coupling constants for two of the $-NH_3$ hydrogen nuclei in good agreement with NQR (nuclear quadrupole resonance) results. We therefore use the ADPs estimated from X-ray data by Destro *et al.* (2000) as reference values for the present analysis.

The general agreement between the present TLS + ONIOM model ADPs and the reference values is reasonable (Fig. 5), although overall agreement ($\bar{S}_{12}$ = 1.11) is markedly worse than found for 1-methyluracil and benzene. It is difficult to conclude whether ADPs derived from the ONIOM calculations are in better agreement with reality than the reference model, but from the results obtained for 1-methyluracil it is likely that the present TLS + ONIOM values are closer to reality, as they include a wider variety of motions. However, we note that the effect of substantial hydrogen bonding in this example may not be adequately modelled by the ONIOM method, even though atomic charges are included in the calculation and intermolecular hydrogen bonding is included (to some extent) in the molecular mechanics force field.

The agreement between the scaled ONIOM normal-mode frequencies and corresponding IR and Raman data shown in Table 3 is quite remarkable. Surprisingly, the frequency of the NH_3 torsion derived from the ONIOM calculation agrees extremely well with the experimental results, and this is one of

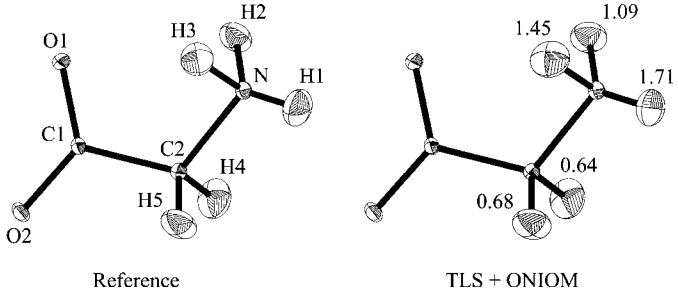

Figure 5
ORTEP representation and atom labelling for α-glycine at 23 K (99% probability level). Reference values are calculated from a mixture of TLS analysis of heavy-atom ADPs, and selected modes and frequencies taken from spectroscopic data (Destro *et al.*, 2000). For the TLS + ONIOM model, S_{12} indices are given for each H atom; $\bar{S}_{12}$ = 1.11.

Table 3
Vibrational frequencies (cm^{-1}) for α-glycine obtained from solid-state IR and Raman spectroscopy (Tsuboi *et al.*, 1958) compared with those from an ONIOM cluster calculation (scaled by 0.90).

Description	IR	Raman	ONIOM
NH_3 torsion	516	–	539
CH_2 rock	910	–	892
NH_3 rock	1115	1115	1115†
CH_2 twist	–	1240	1292
CH_2 wag	1335	1325	1335
CH_2 bend	1450	1440	1433
NH_3 symmetric deformation	1500	–	1557

† Averaged over many similar modes.

the most important vibrational modes, as it is a low-frequency large-amplitude vibration that is expected to be affected by hydrogen bonding. Although the frequencies of these specific modes compare very well, there are also many other low-frequency modes that contribute to the motion of the NH_3 H atoms. The TLS + ONIOM-derived ADPs are in general larger than the ADPs reported by Destro *et al.* (2000) and whether this reflects reality is unknown in the absence of accurate neutron diffraction data at ~20 K. However, we note that the largest discrepancy between our TLS + ONIOM estimates and those of Destro *et al.* is observed for H1, precisely the same H atom for which the agreement between X-ray and NQR quadrupole coupling constants is least favourable (Destro *et al.*, 2000); it would be very interesting to incorporate our TLS + ONIOM ADPs for H atoms in a re-analysis of the 23 K X-ray data.

5.4. Xylitol

Xylitol has been the subject of recent neutron diffraction experiments and charge-density analyses (Madsen *et al.*, 2003, 2004). In that work, ADPs derived from the neutron diffraction data were analysed in terms of the contributions of both the internal and external modes to the total motion. External motion was modelled by a TLS analysis of experimental heavy-atom ADPs and internal contributions to the motion were estimated by two different methods. The first method attempted to estimate the internal contribution to the overall motion *via* an *ab initio* calculation on an isolated molecule of xylitol, but *ab initio* magnitudes of the calculated internal portion of the ADPs for H atoms were larger than the ADPs derived from the neutron diffraction data. This has also been observed on previous occasions for other molecules (Luo *et al.*, 1996) and is due largely to the presence of low-frequency internal modes for the isolated molecule that are non-existent in the presence of a crystalline environment. This is generally not observed for aromatic and small molecules, as they are comparatively rigid. Xylitol, as acknowledged by Madsen *et al.* (2003), requires the crystal environment to be modelled in order for sensible estimates of the internal motion to be derived, and we have remarked earlier on the evidence from the $\Delta_{A,B}$ matrix for significant internal molecular motions.

Owing to the failure of their *ab initio* approach for the determination of sensible internal modes and frequencies, a

Table 4
Average mean-square displacement amplitudes for internal vibrations of H nuclei ($Å^2 \times 10^4$).

The first row lists average values derived from various neutron diffraction experiments (Madsen *et al.*, 2003) and the second row lists the present theoretical estimates based on high-frequency vibrational modes derived from the present ONIOM calculation on xylitol (frequencies scaled by 0.90).

		ε_{11}	ε_{22}	ε_{33}
Methylene (CH_2)	Neutron	51 (11)	145 (33)	246 (72)
	ONIOM	76	163	253
Methine (CH)	Neutron	50 (13)	–	144 (27)†
	ONIOM	73	–	171
Hydroxy (OH)	Neutron	35 (22)	183 (43)	101 (34)
	ONIOM	65	292	160

† It is not possible to unambiguously define an 'in-plane' and 'out-of-plane' bend for methine H atoms, hence the values from Madsen *et al.* (2003) were averaged for comparison.

second approach was adopted by Madsen *et al.* (2003), and involved calculating internal contributions to the overall displacement parameters by assuming them to be the difference between the experimental H-atom ADPs and those inferred from a TLS model based on heavy-atom ADPs. This was reasonably successful, but the rigid-body model was fitted to the ADPs of heavy atoms derived from the neutron experiment, uncorrected for contributions from internal motion; ideally the rigid body should only fit the motion due to the external modes. Hence these mean vibrational amplitudes represent a mixture of internal motion for H atoms and artefacts due to the neglect of the internal motion of the heavy atoms, which may be part of the reason why the e.s.d.s associated with the results (Table 4) are so large. The success of this latter method adopted by Madsen *et al.* (2003) for xylitol may be due partly to the fact that, in their estimation of average mean-square displacement parameters for H atoms, those authors incorporated results obtained for xylitol itself. As a consequence, the experimental neutron data for xylitol is contributing to the approximate ADPs, helping to ensure that the estimated ADPs agree well with the reference neutron results.

Table 4 compares the average internal stretching and bending mean-square amplitudes derived from neutron diffraction data on various compounds (Madsen *et al.*, 2003, 2004) with results from the present ONIOM calculation of xylitol. The amplitudes of the stretching motions (ε_{11}) should be well reproduced by theory regardless of the environment surrounding the molecule, but systematic differences are observed between the ONIOM and averaged experimental values, most likely due to the neglect of corrections for internal motion before determining the external motion model. The ONIOM theoretical mean-square displacement amplitudes for the stretching motions are about 25×10^4 $Å^2$ larger than the corresponding values derived from experiment for all the listed cases. The results in Table 2 show that the agreement for benzene is much better, with the difference being only 6×10^4 $Å^2$, and for this reason we believe that the present ONIOM results are more reliable than the averaged experimental results for stretching motions. For benzene it can also be seen that the bending motions, while reproduced reasonably well, are systematically overestimated, the difference being approximately 20×10^4 $Å^2$ for the out-of-plane vibrations and about half that for the in-plane vibration. From Table 4, the in- and out-of-plane vibrations are reproduced remarkably well for the methylene and methine protons, but the agreement for the hydroxy protons is comparatively poor. This indicates that, despite the effort taken to specify charges and an appropriate surrounding environment in the ONIOM calculations, hydrogen bonding is possibly inadequately modelled in the ONIOM calculation. Despite the problems relating to the overestimate of the in- and out-of-plane bending amplitudes for the hydroxy protons, the ADPs obtained from the TLS + ONIOM model are still in remarkably good agreement with reference ADPs (Fig. 6). The main discrepancies are observed for the hydroxy H atom H13, and the methylene H atoms H5*A* and H5*B*. The reason for this may be the description of hydrogen bonding but, as already discussed above, it is also likely that a rigid body is an inferior description for the heavy-atom skeleton for xylitol; a segmented rigid-body fit may improve this. For comparison with the results in Fig. 6, the best model H-atom ADPs reported by Madsen *et al.* (2004) (labelled TLS:mean in that work) result in $\bar{S}_{12} = 0.55$ for agreement with the original neutron diffraction results.

5.5. 2-Methyl-4-nitroaniline (MNA)

MNA has also been the subject of a recent neutron diffraction experiment and charge-density analysis (Whitten, Turner *et al.*, 2006), with a view to re-assessing the very large dipole moment enhancement in the crystal, which was a major outcome of an earlier charge-density analysis (Howard *et al.*, 1992). The study by Howard *et al.* treated H-atom thermal motion as isotropic, and improved modelling of the motion of

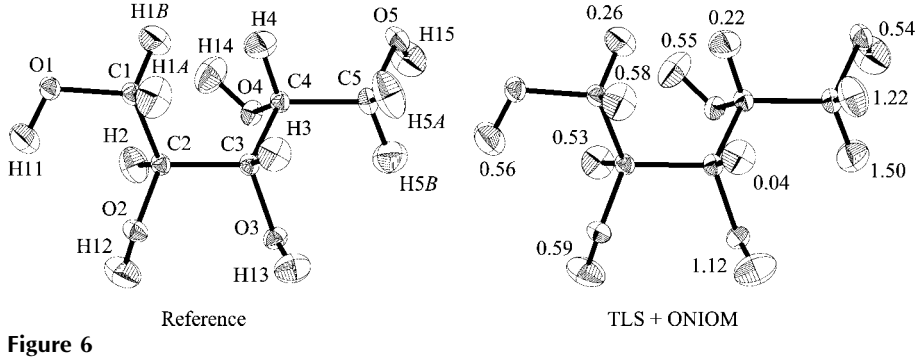

Figure 6
ORTEP representation and atom labelling for xylitol at 123 K (99% probability level). Reference values are calculated from neutron ADPs (Madsen *et al.*, 2003), adjusted against X-ray ADPs (Madsen *et al.*, 2004) using *UIJXN* (Blessing, 1995). For the TLS + ONIOM model, S_{12} indices are given for each H atom; $\bar{S}_{12} = 0.64$.

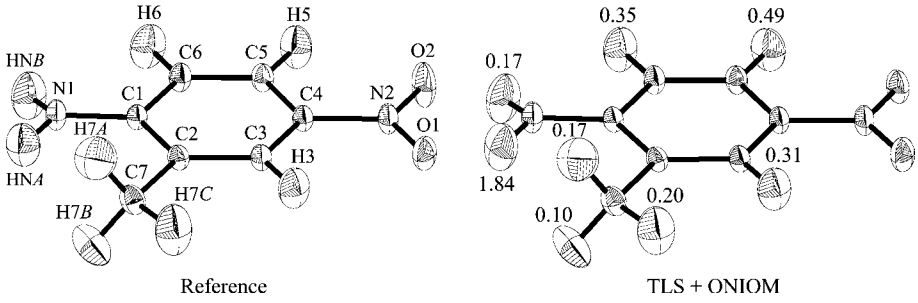

Figure 7
ORTEP representation and atom labelling for MNA at 100 K (99% probability level). Reference values are calculated from neutron ADPs adjusted against X-ray ADPs (Whitten, Turner *et al.*, 2006) using *UIJXN* (Blessing, 1995). For the TLS + ONIOM model, S_{12} indices are given for each H atom; $\bar{S}_{12} = 0.45$.

those atoms was an important motivation for the new experiments. The latest charge-density analysis employed adjusted neutron diffraction ADPs, and we use those as reference values in the present study; Fig. 7 summarizes the comparison between reference and TLS + ONIOM H-atom ADPs. It is clear that for all but one H atom there is remarkably good agreement between the two sets of ADPs, and the overall agreement, $\bar{S}_{12} = 0.45$, is better than that obtained for xylitol, but not as good as for 1-methyluracil or benzene. The only significant discrepancy occurs for HN*A*, which forms by far the shortest H$\cdots$O contact in the crystal (the HN*A*$\cdots$O distance is 2.063 Å, compared with a value of 2.274 Å for HN*B*$\cdots$O), suggesting once again that the description of hydrogen bonding is inadequate in our present ONIOM calculations.

6. Discussion and conclusions

We have shown that a two-layer HF/MM ONIOM method implemented in *GAUSSIAN03* (Frisch *et al.*, 2004) and applied to suitably sized clusters of organic molecules is capable of providing excellent estimates of the internal motion of H atoms in the crystal, and in some ideal cases it can also reproduce reasonable estimates of the external motion. Because of this, the incorporation of this information with estimates of molecular motion from a TLS fit to X-ray ADPs for heavy atoms leads to ADPs for H atoms that are in remarkably good agreement with reference values from neutron diffraction experiments, suitably adjusted. Although results have been presented only for five molecular crystals, we are confident that with some fine-tuning the TLS + ONIOM approach is capable of providing H-atom ADPs to routinely complement modern charge-density studies on organic molecular crystals.

The TLS + ONIOM model explored in some detail in this work involves two key elements: the use of a rigid-body model to fit X-ray ADPs of the heavy-atom molecular skeleton (suitably corrected for contributions from internal modes) to estimate the contribution from external modes, and the use of a two-layer HF/MM ONIOM cluster approach to estimate the contribution to both heavy atoms and H atoms from the internal modes. We have arbitrarily used the $3N - 6$ highest-frequency modes from the ONIOM calculation for this latter purpose, and the indications are that this is both a pragmatic and successful approach, despite the fact that external molecular motions are expected to contribute to many normal modes at higher frequency. However, analysis of X-ray ADPs and TLS + ONIOM results for xylitol also suggest that a segmented rigid-body treatment will most likely be required for molecules with significant internal degrees of freedom. We will explore this in future studies, and note that it has already been employed in a charge-density study of *N*-(4-nitrophenyl)-L-prolinol (Fkyerat *et al.*, 1995), along with estimates of internal C—H vibrations, and the results used in preference to those from a neutron diffraction study at the same temperature.

Bürgi's elegant analysis of multi-temperature neutron diffraction data can also be used to estimate H-atom ADPs (Bürgi *et al.*, 2002), but most of the outcomes from analyses of that kind are of less relevance to the charge-density community, as they pertain to other dynamical aspects of the molecule in the crystal; ADPs of H atoms represent a small subset of this information. The present TLS + ONIOM method is less flexible than the analysis of multi-temperature diffraction data, but is expected to be more accessible to the charge-density community as it does not require additional experimental information beyond that obtained in the X-ray experiment, and the calculations are straightforward to perform and to subsequently analyse. All computations in this work can be performed using *Gaussian03* (Frisch *et al.*, 2004), *THMA*11 (Schomaker & Trueblood, 1998) and *XDVIB* (Koritsanszky *et al.*, 2003); hence, no in-house software is required.

The ONIOM cluster approach, like all methods using a combination of levels of theory, is a trade-off between time and accuracy. The TLS + ONIOM model we have employed makes use of an approximate cluster approach which, while imperfect, is capable of yielding excellent estimates of H-atom ADPs that are otherwise unobtainable. Despite the evident success of the present TLS + ONIOM approach, we believe that there is considerable room for improvement. Some of the results suggest that the description of hydrogen bonding, presently afforded by the molecular mechanics parameters in the UFF force field, is less than satisfactory, and this is hardly surprising. There is scope within the framework of ONIOM calculations in *GAUSSIAN03* to specify link atoms, which are treated differently in the different layers of the calculation, and this may provide an improvement. A simpler option may be to determine a different scale factor for the frequencies of modes that would be affected by hydrogen bonding. However, neither solution is in the spirit of the present work, as the objective was to devise a routine scheme for approximation of H-atom ADPs for charge-density analysis. We expect that it

would be worthwhile exploring the use of other combinations of theory and/or basis sets for high-level and low-level layers, beyond the present HF/MM combination. In this direction, we note that the ONIOM approach has been applied recently to the crystal structure of L-alanine, with a B3LYP/6-31G** description of the central molecule, and semi-empirical PM3 description of the molecules in the low-level layer (Pauwels *et al.*, 2003). Finally, perhaps the most immediate question we have is whether the point charges presently used for electronic embedding in the ONIOM calculations need to come from a crystal Hartree–Fock calculation. There is no doubt that they are appropriate for our purposes and, based on the agreement indices obtained in the fitting process used to determine the charges (Whitten, McKinnon *et al.*, 2006), it is clear that they produce a realistic representation of the crystalline potential. However, we note that periodic *ab initio* calculations are not always routine, especially for larger molecules.

This is our first attempt in this direction, and we have described a number of straightforward ways in which improvements might be made. In the future we anticipate that all the necessary vibrational information may be obtained from periodic *ab initio* calculations. Analytic first and second derivatives are now available from *CRYSTAL03* (Saunders *et al.*, 2003) and there have already been applications to small high-symmetry systems (Pascale *et al.*, 2004); however, the computation time would currently be prohibitive for moderately large systems with low symmetry. Nevertheless, the simpler approach described in this work, based on two-layer ONIOM cluster calculations, clearly represents an important step in the right direction, and we anticipate that incorporation of H-atom ADPs obtained in the manner we have described will go a long way towards overcoming one of the outstanding limitations of current charge-density analyses on organic molecular crystals.

This research is supported by the Australian Research Council. AEW is grateful for the receipt of an Australian Postgraduate Award, as well as a Postgraduate Research Award from the Australian Institute of Nuclear Science and Engineering, which funded his PhD studies. Riccardo Destro and Pietro Roversi are thanked for providing ADPs from their unpublished X-ray study of 1-methyluracil at 21 K; Chris Van Alsenoy for helpful comments regarding some of the *ab initio* calculations; and Parthapratim Munshi for assistance with some of the TLS calculations. We are especially grateful to the referees for their critical and insightful comments.

References

Blessing, R. H. (1995). *Acta Cryst.* B**51**, 816–823.
Bürgi, H. B. (2000). *Annu. Rev. Phys. Chem.* **51**, 275–296.
Bürgi, H. B. & Capelli, S. C. (1999). *Implications of Molecular and Materials Structure for New Technologies*, edited by J. A. K. Howard, F. H. Allen & G. P. Shields, pp. 45–58. Amsterdam: Kluwer Academic.
Bürgi, H. B. & Capelli, S. C. (2000). *Acta Cryst.* A**56**, 403–412.
Bürgi, H. B., Capelli, S. C., Goeta, A. E., Howard, J. A. K., Spackman, M. A. & Yufit, D. S. (2002). *Chem. Eur. J.* **8**, 3512–3521.
Bürgi, H. B. & Förtsch, M. (1999). *J. Mol. Struct.* **485–486**, 457–463.
Bürgi, H. B., Rangavittal, N. & Hauser, J. (2001). *Helv. Chim. Acta*, **84**, 1889–1906.
Capelli, S. C., Förtsch, M. & Bürgi, H. B. (2000). *Acta Cryst.* A**56**, 413–424.
Coppens, P. (2005). *Angew. Chem. Int. Ed.* **44**, 6810–6811.
Coppens, P., Dam, J., Harkema, S., Feil, D., Feld, R., Lehmann, M. S., Goddard, R., Krüger, C., Hellner, E., Johansen, H., Larsen, F. K., Koetzle, T. F., McMullan, R. K., Maslen, E. N. & Stevens, E. D. (1984). *Acta Cryst.* A**40**, 184–195.
Cruickshank, D. W. J. (1956). *Acta Cryst.* **9**, 754.
Cyvin, S. J. (1968). *Molecular Vibrations and Mean Square Amplitudes.* Amsterdam: Elsevier.
Dapprich, S., Komáromi, I., Byun, K. S., Morokuma, K. & Frisch, M. J. (1999). *J. Mol. Struct.* **461–462**, 1–21.
Destro, R., Bianchi, R. & Morosi, G. (1989). *J. Phys. Chem.* **93**, 4447–4457.
Destro, R., Roversi, P., Barzaghi, M. & Marsh, R. E. (2000). *J. Phys. Chem. A*, **104**, 1047–1054.
Dougherty, J. & Spackman, M. A. (1994). *Mol. Phys.* **82**, 193–209.
Dunitz, J. D., Maverick, E. F. & Trueblood, K. N. (1988). *Angew. Chem. Int. Ed.* **27**, 880–895.
Dunitz, J. D., Schomaker, V. & Trueblood, K. N. (1988). *J. Phys. Chem.* **92**, 856–867.
Fkyerat, A., Guelzim, A., Baert, F., Paulus, W., Heger, G., Zyss, J. & Périgaud, A. (1995). *Acta Cryst.* B**51**, 197–209.
Flaig, R., Koritsanszky, T., Zobel, D. & Luger, P. (1998). *J. Am. Chem. Soc.* **120**, 2227–2238.
Frisch, M. J. *et al.* (2004). *GAUSSIAN03.* Revision C.02. Gaussian, Wallingford, CT, USA.
He, X.-M. & Craven, B. M. (1993). *Acta Cryst.* A**49**, 10–22.
Hirshfeld, F. L. (1976). *Acta Cryst.* A**32**, 239–244.
Howard, S. T., Hursthouse, M. B., Lehmann, C. W., Mallinson, P. R. & Frampton, C. S. (1992). *J. Chem. Phys.* **97**, 5616–5630.
Jeffrey, G. A., Ruble, J. R., McMullan, R. K. & Pople, J. A. (1987). *Proc. R. Soc. London Ser. A*, **414**, 47–57.
Jonsson, P.-G. & Kvick, Å. (1972). *Acta Cryst.* B**28**, 1827–1833.
Klooster, W. T., Swaminathan, S., Nanni, R. & Craven, B. M. (1992). *Acta Cryst.* B**48**, 217–227.
Koritsanszky, T., Howard, S. T., Macchi, P., Gatti, C., Farrugia, L. J., Mallinson, P. R., Volkov, A., Su, Z., Richter, T. & Hansen, N. K. (2003). *XD – A Computer Program Package for Multipole Refinement and Analysis of Charge Densities from Diffraction Data*, 4.10 ed. Freie Universität Berlin, Germany.
Koritsanszky, T. S. & Coppens, P. (2001). *Chem. Rev.* **101**, 1583–1627.
Lenstra, A. T. H., Van Loock, J. F. J., Rousseau, B. & Maes, S. T. (2001). *Acta Cryst.* A**57**, 629–641.
Lewis, T. P., Miles, H. T. & Becker, E. D. (1984). *J. Phys. Chem.* **88**, 3253–3260.
Luger, P., Wagner, A., Hübschle, C. B. & Troyanov, S. I. (2005). *J. Phys. Chem. A*, **109**, 10177–10179.
Luo, J.-Q., Ruble, J. R., Craven, B. M. & McMullan, R. K. (1996). *Acta Cryst.* B**52**, 357–368.
Machida, K., Kagayama, A., Saito, Y., Kuroda, Y. & Uno, T. (1977). *Spectrochim. Acta A*, **33**, 569–574.
Madsen, A. O., Mason, S. & Larsen, S. (2003). *Acta Cryst.* B**59**, 653–663.
Madsen, A. O., Sorensen, H. O., Flensburg, C., Stewart, R. F. & Larsen, S. (2004). *Acta Cryst.* A**60**, 550–561.
Maseras, F. & Morokuma, K. (1995). *J. Comput. Chem.* **16**, 1170–1179.
May, E., Destro, R. & Gatti, C. (2001). *J. Am. Chem. Soc.* **123**, 12248–12254.
McMullan, R. K. & Craven, B. M. (1989). *Acta Cryst.* B**45**, 270–276.
McMullan, R. K., Koetzle, T. F. & Fritchie, C. J. (1997). *Acta Cryst.* B**53**, 645–653.
Morokuma, K. (2003). *Bull. Korean Chem. Soc.* **24**, 797–801.
Munshi, P. & Guru Row, T. N. (2005). *Crystallogr. Rev.* **11**, 199–241.

Ochterski, J. W. (1999). *Vibrational Analysis in Gaussian*, http://www.Gaussian.com/g_whitepap/vib.htm.
Oddershede, J. & Larsen, S. (2004). *J. Phys. Chem. A*, **108**, 1057–1063.
Pascale, F., Zicovich-Wilson, C. M., Lopez Gejo, F., Civalleri, B., Orlando, R. & Dovesi, R. (2004). *J. Comput. Chem.* **25**, 888–897.
Pauwels, E., Van Speybroeck, V. & Waroquier, M. (2003). *Int. J. Quantum Chem.* **91**, 511–516.
Rappé, A. K., Casewit, C. J., Colewell, K. S., Goddard, W. A. & Skiff, W. M. (1992). *J. Am. Chem. Soc.* **114**, 10024–10035.
Rosenfield, R. E., Trueblood, K. N. & Dunitz, J. D. (1978). *Acta Cryst.* A**34**, 828–829.
Rousseau, B., Keulers, R., Dessyn, H. O., Geize, H. J. & Van Alsenoy, C. (1999). *Chem. Phys. Lett.* **302**, 55–59.
Rousseau, B., Maes, S. T. & Lenstra, A. T. H. (2000). *Acta Cryst.* A**56**, 300–307.
Rousseau, B., Van Alsenoy, C., Keulers, R. & Dessyn, H. O. (1998). *J. Phys. Chem. A*, **102**, 6540–6548.
Roversi, P., Barzaghi, M., Merati, F. & Destro, R. (1996). *Can. J. Chem.* **74**, 1145–1161.
Roversi, P. & Destro, R. (2004). *Chem. Phys. Lett.* **386**, 472–478.
Saunders, V. R., Dovesi, R., Roetti, C., Orlando, R., Zicovich-Wilson, C. M., Harrison, N. M., Doll, K., Civalleri, B., Bush, I. J., D'Arco, P. & Llunell, M. (2003). *CRYSTAL*03 *User's Manual*, 1.0 ed. University of Torino, Italy.
Schomaker, V. & Trueblood, K. N. (1968). *Acta Cryst.* B**24**, 63–76.
Schomaker, V. & Trueblood, K. N. (1998). *Acta Cryst.* B**54**, 507–514.
Scott, A. P. & Radom, L. (1996). *J. Phys. Chem.* **100**, 16502–16513.
Svensson, M., Humbel, S., Froese, R. D. J., Matsubara, T., Sieber, S. & Morokuma, K. (1996). *J. Phys. Chem.* **100**, 19357–19363.
Szczesniak, M., Nowak, M. J., Szczepaniak, K. & Person, W. B. (1985). *Spectrochim. Acta*, **41**, 237–250.
Thakkar, A. J., Koga, T., Saito, M. & Hoffmeyer, R. E. (1993). *Int. J. Quantum Chem. Quantum Chem. Symp.* **27**, 343–354.
Tsuboi, M., Onishi, I., Nakagawa, I., Shimanouchi, T. & Mizushima, S.-I. (1958). *Spectrochim. Acta*, **12**, 253–261.
Vreven, T., Byun, K. S., Komáromi, I., Dapprich, S., Montgomery, J. A., Morokuma, K. & Frisch, M. J. (2006). *J. Chem. Theor. Comput.* **2**, 815–826.
Vreven, T., Morokuma, K., Farkas, O., Schlegel, H. B. & Frisch, M. J. (2003). *J. Comput. Chem.* **24**, 760–769.
Weber, H. P., Craven, B. M., Sawzik, P. & McMullan, R. K. (1991). *Acta Cryst.* B**47**, 116–127.
Whitten, A. E., McKinnon, J. J. & Spackman, M. A. (2006). *J. Comput. Chem.* **27**, 1063–1070.
Whitten, A. E., Turner, P., Klooster, W. T., Piltz, R. O. & Spackman, M. A. (2006). *J. Phys. Chem. A.* **110**, 8763–8776.
Willis, B. T. M. & Pryor, A. W. (1975). *Thermal Vibrations in Crystallography*. Cambridge University Press.
Zheng, A., Yang, M., Yue, Y., Ye, C. & Deng, F. (2004). *Chem. Phys. Lett.* **399**, 172–176.