FOUNDATIONS
ADVANCES

ISSN: 2053-2733
journals.iucr.org/a

An improved experimental databank of transferable multipolar atom models – ELMAM2. Construction details and applications

Sławomir Domagała, Bertrand Fournier, Dorothee Liebschner, Benoît Guillot and Christian Jelsch

Acta Cryst. (2012). A**68**, 337–351

Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 4 November 2011
Accepted 23 February 2012

An improved experimental databank of transferable multipolar atom models – ELMAM2. Construction details and applications

Sławomir Domagała,[a,b] Bertrand Fournier,[a] Dorothee Liebschner,[a] Benoît Guillot[a] and Christian Jelsch[a]*

[a]Laboratoire de Cristallographie, Résonance Magnétique et Modélisations (CRM2), CNRS, UMR 7036, Institut Jean Barriol, Faculté des Sciences et Technologies, Nancy Université, BP 70239, 54506 Vandoeuvre-lès-Nancy Cedex, France, and [b]Department of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warsaw, Poland. Correspondence e-mail: christian.jelsch@crm2.uhp-nancy.fr

ELMAM2 is a generalized and improved library of experimentally derived multipolar atom types. The previously published ELMAM database is restricted mostly to protein atoms. The current database is extended to common functional groups encountered in organic molecules and is based on optimized local axes systems taking into account the local pseudosymmetry of the molecular fragment. In this approach, the symmetry-restricted multipoles have zero populations, while others take generally significant values. The various applications of the database are described. The deformation electron densities, electrostatic potentials and interaction energies calculated for several tripeptides and aromatic molecules are calculated using ELMAM2 electron-density parameters and compared with the former ELMAM database and density functional theory calculations.

1. Introduction

Ultra-high-resolution X-ray diffraction crystallography is a unique technique that allows one to obtain the experimental distribution of the electron density in crystals. Charge-density determination is now a mature branch of modern crystallography with many publications in a variety of journals, focusing on an increasing range of inorganic, organometallic, organic and biological materials (Coppens, 1997; Spackman, 1997; Koritsanszky & Coppens, 2001; Munshi & Guru Row, 2005*a*). The electron-density distribution is frequently modelled *via* the Hansen & Coppens multipolar model (Hansen & Coppens, 1978), where the individual atomic densities are described in terms of spherical core and valence densities with an expansion of atom-centred real spherical harmonic functions. Thus experimentally derived densities can be compared with some success with the charge densities obtained from high-level theoretical calculations, despite the experimental errors and the approximations used in the multipolar expansion model. A range of problems of chemical and physical interest have been successfully resolved using this technique (Spackman, 1992; Coppens, 1997; Tsirelson & Ozerov, 1996). Unfortunately, because of the high demands of the crystal quality and measurement conditions, the number of publications involving new high-resolution structural studies is rather limited. Owing to these constraints, the idea of constructing the experimentally derived charge density has emerged (Brock *et al.*, 1991). This idea was quickly noticed and the first database of experimentally obtained charge-density parameters was constructed, ELMAM (Pichon-Pesme *et al.*, 1995). Other initiatives have been undertaken to construct such libraries from quantum-mechanical computations of selected small molecules. Two such libraries were constructed: the University at Buffalo Pseudoatom Databank (UBDB; Volkov *et al.*, 2004; Dominiak *et al.*, 2007) and the Invariom database (Dittrich *et al.*, 2004; Dittrich, Hübschle *et al.*, 2006). Improvements to the residual electron density, geometric parameters and atomic displacement parameters when using database electron-density parameters have been thoroughly discussed (Jelsch *et al.*, 1998, 2005; Dittrich *et al.*, 2005, 2007, 2008; Dittrich, Hübschle *et al.*, 2006, 2009; Dittrich, Strümpel *et al.*, 2006; Dittrich, Weber *et al.*, 2009; Volkov *et al.*, 2007; Zarychta *et al.*, 2007; Bąk *et al.*, 2009). The potential applications of the databases to macromolecules (Muzet *et al.*, 2003; Guillot *et al.*, 2008) and in the computations of electrostatic interaction energies between host–guest protein complexes were also investigated (Dominiak *et al.*, 2009; Fournier *et al.*, 2009).

The ELMAM database has been extended from protein atom types to common organic molecules. New chemical environments (atom types) can be easily added to the database when new charge-density diffraction data become publicly available. The improved database hereafter referred to as ELMAM2 is based on the optimal local coordinate

systems (Domagała & Jelsch, 2008). The aim of the current work is to present a thorough comparison of the improved database with the previous experimental database as well as with the AMBER point charges databank (Case *et al.*, 2008) and theoretically obtained electron-density distributions. The comparison involves all the frequently studied features and derived properties of the electron-density distribution. A detailed comparison of the theoretical and experimental databases of the aspherical atom types was presented by Bąk *et al.* (2011). A typical example of applications of the new ELMAM2 database for the common aromatic systems is given in the study of quercetin monohydrate (Domagała *et al.*, 2011). The databank transfer procedure can be conveniently applied to crystal structures of small molecules at usual resolution to yield a more accurate structure and better crystallographic statistics (Ahmed *et al.*, 2011).

2. Construction of the generalized ELMAM2 database

2.1. Multipolar refinements of the selected structures

A set of 54 high-resolution structures was selected for the construction of the generalized Experimental Library of Multipolar Atom Model, hereafter called ELMAM2. The list of all selected structures is given in Table 1S (the 'S' signifies supplementary table or figure);[1] they are taken from the following references: Benabicha *et al.* (2000); Birkedal *et al.* (2004); Bouhmaida *et al.* (2009); Chen *et al.* (2007); Coppens *et al.* (1999); Dahaoui *et al.* (1999); Destro *et al.* (1988); Dittrich *et al.* (2002, 2007); Dittrich & Spackman (2007); Domagała *et al.* (2009); Dominiak *et al.* (2003); Espinosa *et al.* (1996); Fournier *et al.* (2009); Ghermani *et al.* (2004); Guillot *et al.* (2003); Howard *et al.* (2009); Hübschle *et al.* (2008); Kalinowski *et al.* (2007); Luger *et al.* (2004); Lutz *et al.* (2008); Madsen *et al.* (2004); Martin & Pinkerton (1998); Meents *et al.* (2008); Munshi & Guru Row (2002, 2005*b*; 2006*a*,*b*); Munshi *et al.* (2006); Ogawa *et al.* (2006); Overgaard & Hibbs (2004); Parrish *et al.* (2006); Pichon-Pesme *et al.* (2000); Rodrigues *et al.* (2001); Scheins *et al.* (2004); Slouf *et al.* (2002); Sørensen *et al.* (2003); Sparkes *et al.* (2008); Volkov *et al.* (2000); Wiest *et al.* (1994); Zhurov *et al.* (2005); Zhurova *et al.* (2002, 2006); Zhurova & Pinkerton (2001). The high-resolution structures available in the literature were surveyed in order to find accurate electron-density determinations. Their coordinates and structure factors were obtained either from the journal website (IUCr publications) or directly from the authors. All structures found were refined using our standard strategy. The 'good' data were selected for further averaging. Several factors were considered to qualify the data, *e.g.* featureless residual density maps, reasonable multipolar parameters and no convergence problems. Coordinated metal atoms were not analysed and were not included in the present database.

The charge-density least-squares refinement program *MoPro* (Guillot *et al.*, 2001; Jelsch *et al.*, 2005) was used to perform the multipolar refinements of the selected compounds, applying a standardized refinement strategy.

The atom model used for the description of the total molecular electron density is based on the Hansen & Coppens (1978) multipole formalism. The individual atomic densities are described in terms of spherical core and valence densities with an expansion of atom-centred real spherical harmonic functions. The total atomic electron density is therefore a sum of three components:

$$\rho_{\rm atom}(\mathbf{r}) = \rho_{\rm core}(\mathbf{r}) + P_{\rm val}\kappa^3\rho_{\rm val}(\kappa\mathbf{r}) + \sum_{l=0}^{l_{\max}} \kappa_l'^3 R_{nl}(\kappa'\mathbf{r}) \sum_{m=0}^{l} P_{lm\pm}\, y_{lm\pm}(\theta,\varphi), \qquad (1)$$

where $\rho_{\rm core}$ and $\rho_{\rm val}$ are spherical core and valence densities, respectively. The third term contains the sum of the angular functions $y_{lm\pm}(\theta, \varphi)$ to take into account aspherical deformations. The angular functions $y_{lm\pm}(\theta, \varphi)$ are real spherical harmonic functions. The coefficients $P_{\rm val}$ and $P_{lm\pm}$ are multipole populations for the spherical valence and multipolar density, respectively. The κ and κ' are scaling parameters, which determine the expansion/contraction of the spherical and multipolar valence densities, respectively. In the Hansen–Coppens (Hansen & Coppens, 1978) formalism, the $P_{\rm val}$, $P_{lm\pm}$, κ and κ' are refined parameters together with the atomic coordinates (noted as xyz) and atomic displacement parameters (ADPs).

Several authors have described the general strategies of multipolar refinement (Hansen & Coppens, 1978; Coppens, 1997; Hoser *et al.*, 2009). We also developed a suitable approach for building the generalized multipolar database, based on the high-order refinement of xyz and U_{ij} for the non-H atoms and low-order refinement for the H atoms. The multipolar populations are refined in a stepwise manner, *i.e.* they are added consecutively to the whole set of refined parameters. The non-H atoms are refined up to octupolar level with the exception of heavy atoms ($Z > 10$), which are refined up to hexadecapolar level. For the H atoms, only one bond-oriented dipole and one quadrupolar function (q_{3z^2-1}) are used. The *SHADE* program (Simple Hydrogen Anisotropic Displacement Estimator; Madsen, 2006) was used to estimate ADPs (noted as U_{ij}) for H atoms. During the refinement, numerous restraints and/or constraints are used to avoid unphysical values of the refined parameters. The detailed refinement procedure is summarized in the following steps:

(i) The X—H distances are restrained in the subsequent refinements to the standard neutron distance (Allen *et al.*, 1987, 2006) with a restraint σ = 0.002 Å. Restraints are also applied to the isotropic thermal displacement parameters (noted as $U_{\rm iso}$) of H atoms, $U_{\rm iso} = kU_{\rm eq}(X)$, where k = 1.2, except for –CH_3, –NH_3^+, –OH and –SH groups for which k was set to 1.5 as they have a rotational degree of freedom.

(ii) High-order refinement ($\sin\theta/\lambda > 0.7$ Å^{-1}) of the xyz and U_{ij} parameters for the non-H atoms. Then, considering the whole resolution range, refinement of the scale factor, xyz and $U_{\rm iso}$ parameters of H atoms. These steps were repeated until convergence was obtained.

[1] Supplementary material for this paper is available from the IUCr electronic archives (Reference: PC5007). Services for accessing this material are described at the back of the journal.

(iii) The anisotropic ADPs of the H atoms were estimated using the *SHADE* program and were kept strongly restrained to the *SHADE* values in the subsequent refinements.

(iv) (*a*) Block refinement against high-order reflections of xyz, U_{ij} for non-H atoms. Block refinement against all reflections of: (*b*) scale factor; (*c*) xyz and U_{ij} for H atoms; (*d*) $P_{\rm val}$; (*e*) P_{lm}; (*f*) κ; (*g*) κ'.

The charge-density parameters ($P_{\rm val}$, P_{lm}, κ, κ') were introduced in the refinement step by step, repeating every step until convergence was obtained.

During the refinement, κ' parameters of H atoms were strongly restrained (σ_R = 0.01) to theoretical values which depend on the type of the carrying atom to which the H atom is attached (Volkov *et al.*, 2001). For example, the κ' parameters were restrained to 1.18, 1.40 and 1.50 for H atoms connected to C, N and O atoms, respectively. Restraints were applied on (κ, $P_{\rm val}$) parameters to preserve a linear relation between the expansion–contraction of the valence shell (κ) and the net atomic charge $q = P_{\rm neut} - P_{\rm val}$ (Volkov *et al.*, 2001; Jelsch *et al.*, 2005).

2.1.1. Constraints, restraints and atom-type selection rules. During the multipolar refinements, atoms within a similar chemical environment had their charge-density parameters $P_{\rm val}$, P_{lm}, κ and κ' values constrained to be identical. Moreover, optimal local axes were used and the highest possible symmetry was imposed on the refined multipoles (Domagała & Jelsch, 2008). An algorithm built into the *MoPro* program generates automatically the multipolar local axes systems of all the atoms of the molecule in a unique manner.

At first, a connectivity list is generated for all the atoms. For unique ordering of the neighbours of a central atom, the list of considered atoms is sorted according to the following criteria of decreasing importance: (*a*) decreasing atomic numbers, (*b*) decreasing number of bonds, (*c*) decreasing atomic numbers of neighbours, (*d*) increasing distances to the central atom.

The type of local axes system is assigned to an atom according to the number of neighbours and its symmetry. The multipolar refinement, as described before, is performed and the final multipolar parameters, including estimated standard deviations (e.s.d.'s), are stored in the output file.

2.2. Averaging atoms from different molecules

The electron-density parameters resulting for each molecule from its final multipolar refinement were used as an input to a Perl script to create the atom types stored further in the database. Atom-selection procedures and averaging of atom types from different molecules were carried out. Non-H atoms with kappa parameters outside a well defined range ($0.9 < \kappa < 1.1$ and $0.7 < \kappa' < 1.3$) were automatically removed from the atom list. This was done to check the correctness of the electron-density parameters and also to ensure that all atoms are coherent with the standard radial function parameters used in the refinements. When the charge-density parameters of several atoms in a given molecule were constrained together by chemical equivalence constraints, only one reference atom was retained in the list of atoms contributing to an atom type.

The chemical surrounding of a given atom type can be represented as a graph. The nodes denote the atoms which are in the vicinity of the considered atom. The atoms connected directly to the considered atoms are called the first-shell neighbours. The atoms connected to the first-shell atoms are called the second-shell neighbours *etc.* The connectivity graphs were represented as strings to facilitate the comparison of atom types. An example of graph and string representation is shown in Fig. 1. The distinction of the different shells is made according to the different parentheses.

Atoms were assigned to the same atom type when the following parameters were exactly the same: chemical type, type of local axes, number and type of neighbours (*first shell*), planarity. Distances and angles between the considered atom and the first neighbours were compared within a certain tolerance (0.05 Å and 5°). A new atom type was created in the databank every time the program encountered an atom with different parameters. In total, the 68 different atom types were encountered when parsing the multipolar parameters of the refined structures.

For most of the atom types, only the first shell of neighbours was exactly compared. These include C, S, F and N atoms. The *second-* and the *third-shell* neighbours were also compared for certain cases of atom types. For example, in the case of H atoms, the *second-shell* neighbours were also taken into account, whereas oxygen, with its many chemical properties, is an atom type which is sensitive to the type of *third-shell* atoms.

This was the case, for instance, to distinguish between carboxylate Oc[o**(**]c(x)], carboxylic Oc[o**(h)**c(x)] and ester Oc[o**(c)**c(x)] oxygen atoms, as the values of multipoles and the chemical properties are different for these atoms. The differences in the third shell are here marked using underlined bold typeface. The electron-density parameters for each atom type were represented as the weighted mean over all atoms contributing to one atom type. The weights were set recipro-

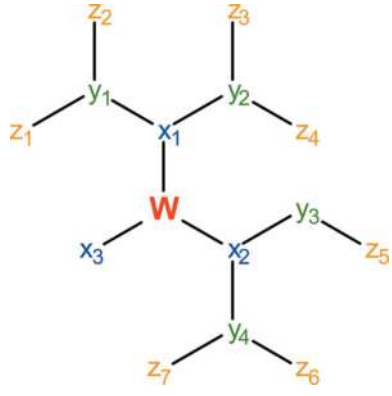

Figure 1
Graph representation of the following chemical string $Wx_1[y_1(z_1z_2)$-$y_2(z_3z_4)]x_2[y_3(z_5)y_4(z_6z_7)]x_3$. W denotes the main atom. The x, y and z atoms correspond to neighbours in the first, second and third shell, respectively. In the case of the string representation, the beginning and the end of each shell is denoted by parentheses, if it is necessary to do so. Therefore, the second shell is distinguished by the '[' and ']' characters, whereas the third one is differentiated by the '(' and ')' characters.

cally proportional to the squared e.s.d.'s of the refined multipolar and kappa parameters. When several atoms in one molecule were chemically equivalent, their common charge-density parameters were constrained to be identical. This resulted in decreased e.s.d. values and as a consequence increased the weight of the atom type in the averaging procedure.

The average values over all other parameters (*e.g.* distances, angles) were computed as the non-weighted arithmetic mean. The resulting average electron-density parameters of the different atom types with the concomitant geometric parameters were saved in a formatted database file. In total 68 atom types were created in the present version. Table 2S contains a description of the selected atom types currently available in the ELMAM2 database.

2.3. Automatic database transfer

The automatic charge-density transfer from the ELMAM2 database can be performed on organic molecules and biological macromolecules, using the *MoPro* program. During the transfer procedure, the atoms in the molecule are compared with the atom types stored in the databank. When a matching atom type is found, the corresponding multipolar parameters (including κ and κ') and the local axes system are transferred to the selected atom. The comparison is made using the same parameters and procedures as for the creation of the databank. No modification is made to atoms which are not identified by the matching procedure. They remain spherical and neutral. However, the user can select an atom type with similar chemical surrounding and impose the transfer of multipolar parameters.

The total charge of the molecule after transfer, which is calculated from the $P_{\rm val}$ monopole populations, can be corrected to the selected formal charge (usually to that of the neutral molecule). The monopole population shifts are performed according to the Faerman & Price (1990) procedure, which takes into account the e.s.d.'s of monopole populations.

3. Methods and applications

3.1. Theoretical calculations

Periodic quantum-mechanical calculations using *CRYSTAL06* (Dovesi *et al.*, 2008) were performed on a selection of crystal structures for comparison and validation of the experimental database.

Known structures of tripeptides were selected to test the reliability of the database approach for biological compounds and to compare the quality of the transferred electron density with respect to the ELMAM and ELMAM2 databases of experimental multipolar atoms. The following tripeptides were analysed: AlaGlyAla anhydrous (Padiyar & Seshadri, 1996), AlaGlyAla monohydrate (Förster *et al.*, 2005), AlaProAla monohydrate (Kalinowski *et al.*, 2007), AlaTyrAla ethanol (Chęcińska *et al.*, 2006) and one structure of glycine (Destro *et al.*, 2000). In order to test the reliability of the generalized database for common chemical compounds some aromatic compounds have been selected.

The following aromatic compounds were used for the examination: benzene (Bacon *et al.*, 1964), catechol (Fronczek *et al.*, 2002), resorcinol (Bacon & Jude, 1973), *p*-nitrophenol forms α and β (Kulkarni *et al.*, 1998), *p*-nitroaniline (Nieger, 2007), *p*-nitrobenzoic acid (Groth, 1980), 2,5-dihydroxybenzoic acid (Cohen *et al.*, 2007), *p*-dinitrobenzene (Tonogaki *et al.*, 1993) and quercetin monohydrate (Domagała *et al.*, 2011). The starting coordinates for all the molecules were taken from the reported crystal structures. The X—H distances were shifted towards average neutron distances (Allen *et al.*, 1987, 2006) for the tripeptide structures using the *MoPro* program. The high-resolution structure of glycine (Destro *et al.*, 2000) was not modified, as the atomic coordinates were taken from the high-resolution structure with the position of the H atoms already corrected. The following approach was adopted for the aromatic compounds. The neutron or high-resolution structures were not modified; the remaining structures had their H-atom positions optimized (see Table 3S for details). The calculations were performed using the obtained geometries with the density functional theory (DFT) method (Hohenberg & Kohn, 1964) and with the B3LYP hybrid functional (Lee *et al.*, 1988; Becke, 1993) using the 6-31G(d,p) basis set (Hariharan & Pople, 1973). The index generation scheme proposed by Le Page & Gabe (1979) was applied to generate unique Miller indices up to s = 1.25 Å^{-1} reciprocal resolution for each structure. Option XFAC of the *CRYSTAL06* program was then used to generate a set of theoretical structure factors from the computed electron densities and using sets of prepared indices.

3.2. Theoretical multipole modelling

The *MoPro* package was used to perform the multipolar refinements (based on F with unitary weighting scheme) against the whole set of generated theoretical structure factors for the selected crystal structures. The corresponding models are referred to as THEO. The non-H atoms were modelled up to the hexadecapolar level and H atoms refined with one dipole d_z and one quadrupole q_{3z^2-1} component. Only valence and multipole populations, κ and κ' parameters were refined while the atomic positions were kept fixed. To consider a static model, the U_{ij} tensor elements were set to zero. One scale-factor parameter was refined. No restraints/constraints were imposed on any atoms, except chemical equivalence constraints on kappa parameters of H atoms attached to the same atoms and connected to similar groups. In the structures containing solvent molecules, no charge transfer was allowed between the different molecules in order to keep them neutral and to allow better comparison with the database-transferred models.

3.3. Database-transferred models

The charge-density parameters were transferred from the ELMAM and ELMAM2 databases to the structures of

selected peptides and aromatic compounds using always the molecular geometry obtained from THEO models (see §3.2).

The automatic transfer procedure of the ELMAM protein charge-density database, as built in *MoPro*, is based on the atoms' names nomenclature of amino acids and applied to the structures of glycine and the four tripeptides. The multipolar parameters from the former ELMAM database were also transferred manually to some of the aromatic molecules (except nitroaniline and quercetin), taking the values from approximately similar atom types in the tyrosine residue and nitro groups in the ELMAM database. For full information about the database transfer, see Table 3S.

An automatic procedure, based on the chemical connectivity recognition described in §§2.2 and 2.3, was used to transfer the proper multipolar values to all the crystal structures from the ELMAM2 generalized database.

Point charges were also compared with the THEO, ELMAM and ELMAM2 multipolar modelling. For the tripeptides, the point charges of the AMBER database, version 10 (Case *et al.*, 2008), were transferred from the library, based on the amino acids' standard nomenclature and depending on the location of the considered atom in the tripeptide (either from the C terminus, N terminus or from the bulk amino acid). In the case of glycine, the point charges values of the C_α, $H_{\alpha1}$ and $H_{\alpha2}$ atoms were taken as the average of the corresponding atomic charges in bulk glycine, C-terminal and N-terminal glycine existing in AMBER.

A charge-scaling procedure (Faerman & Price, 1990) was applied to all the structures to neutralize the total molecular charge after transfer, with unitary weighting (ELMAM) and reciprocally proportional to variances of the $P_{\rm val}$ values (ELMAM2). The AMBER charges of the amino acids in peptides have already formal values of 0, +1 or −1 (they depend on the position on the polypeptide: standard, N-terminus, C-terminus). Therefore, the zwitterionic tripeptides were already neutral with the AMBER point charges. Glycine, however, needed an additional neutralization; the procedure used was the same as for ELMAM.

After the different transfers, the charge-density distribution was described by point charges only (AMBER), and multipolar expansion up to octupolar level for the heavy atoms and bond-directed dipolar level for the H atoms for ELMAM. The ELMAM2 description was similar to ELMAM, except for H atoms, which were quadrupolar. In the THEO multipolar refinement, heavy atoms were modelled up to the hexadecapolar level; for H atoms one bond-directed dipole and one quadrupole were used.

3.4. Monopole, dipole and quadrupole moments calculations

Monopole charges, dipole and quadrupole moments were derived from the multipole populations using the formulas given in Spackman (1992), Coppens (1997) and calculated using the *VMoPro* module, a part of the *MoPro* package. The first and second moments of the charge distribution were computed with respect to the centre of mass of the considered molecule. The traceless convention was used for the quadrupole moment and eigenvalues (Q_{xx}, Q_{yy}, Q_{zz}) were determined.

3.5. AIM charges calculations

The atoms in molecules (AIM) topological charges were computed from the total electron density using three-dimensional grids (0.03 Å grid intervals) using the Bader charge-analysis program (Tang *et al.*, 2009). The grids contained contributions of the parent molecules and of the nearest neighbours to include the crystal environment effect. All the grids were produced with *VMoPro*, using the final molecular models. The AIM charges were computed only for the multipolar-atom-based models.

3.6. Electrostatic interaction energy

The electrostatic interaction energy between two molecules *A* and *B* was computed with *VMoPro* as an integration over the space of the electron density of molecule *A* multiplied by the electrostatic potential generated by molecule *B*, or reciprocally:

$$E_{\rm elec} = \int \rho_A \varphi_B \, {\rm d}r_A = \int \rho_B \varphi_A \, {\rm d}r_B. \qquad (2)$$

Integrals were computed using a numerical integration method based on a spherical grid around selected atoms. The Gauss–Chebyshev (Becke, 1988) and Lebedev & Laikov (1999) quadratures were used for the radial and angular parts, respectively. Radial coordinates and weights were remapped using the formula of Treutler & Ahlrichs (1995). The integrations involved 100 radial and 434 angular quadrature points. Interaction energies were calculated between pairs of neighbouring molecules in contact in the crystal for which two atoms were separated by a distance lower or equal to the sum of their van der Waals radii.

3.7. Grid properties calculations

Statistical analyses were performed on three-dimensional grids of the deformation electron density on the asymmetric unit with a 0.1 Å sampling step. The electrostatic potential (ESP) $V(\mathbf{r}_i)$ was computed in a grid with a 0.1 Å sampling step around the whole molecule with a 3 Å margin. The calculation was performed using all the grid points for the deformation density. The ESP statistics were evaluated around the van der Waals surface within a shell of 0.3 Å thickness. In order to quantify better the ESP distribution, the ESP surface quantities were calculated as proposed by Politzer and co-workers (Murray & Politzer, 1998; Murray *et al.*, 2000). All the notations used here to describe the quantities are from their original papers.

3.8. Refinement of the SerVal peptide

The L-seryl–L-valine crystal structure (Moen *et al.*, 2004) was selected to show the improvement of the database-transferred models over the independent-atom model (IAM). The crystal structure was imported to *MoPro* format. The *X*—H distances were elongated to the standard neutron distances (Allen *et al.*, 1987, 2006). Initially the SerVal structure was

Table 1
Refinement statistics for the L-Ser–L-Val structure.

All the *R*-factor values are in % and residual densities are in e Å^{-3}.

Model	$R(F)$	$wR^2(F)$	$\Delta\rho_{min}$	$\Delta\rho_{max}$
IAM_U_{iso}	3.525	4.180	−0.294	0.266
IAM_U_{anis}	3.606	4.295	−0.300	0.296
ELMAM_U_{iso}	2.267	2.484	−0.167	0.135
ELMAM_U_{anis}	2.290	2.507	−0.176	0.141
ELMAM2_U_{iso}	2.297	2.548	−0.164	0.130
ELMAM2_U_{anis}	2.220	2.421	−0.166	0.129

Table 2
Refinement statistics for the AlaProAla tripeptide.

All the *R*-factor values are in % and residual densities are in e Å^{-3}. The average error of the electron density, $\langle\sigma(\Delta\rho)\rangle$ (Rees, 1976), is in e Å^{-3}.

	ELMAM	ELMAM2	EXP
$wR^2(F)$	2.207	1.968	1.449
$\Delta\rho_{min}$	−0.474	−0.302	−0.246
$\Delta\rho_{max}$	0.334	0.356	0.226
RMS($\Delta\rho$)	0.048	0.046	0.042
$\langle\sigma(\Delta\rho)\rangle$	0.036	0.034	0.031

modelled using the IAM approximation. Atomic displacement parameters, positions (*xyz* coordinates) and the scale factor were refined using the reflections weighting scheme $w(F_{hkl}) = 1/\sigma_F^2$ against the whole diffraction data set, which was measured up to $s = 0.64$ Å^{-1}. The full-matrix least-squares refinement was based on *F*. During the refinement the *X*—H distances were held restrained and the ADPs of the H atoms were scaled according to the U_{eq} of the carrying atom in an analogous way to *SHELXL* (Sheldrick, 2008). Thus, at the end of the refinement, the IAM_U_{iso} model was obtained. Then, the charge-density parameters were transferred from the ELMAM and ELMAM2 databases to the IAM_U_{iso} model. Further refinement of the scale factor, atomic fractional coordinates and atomic thermal parameters was performed until convergence was reached and the following models were stored: ELMAM_U_{iso} and ELMAM2_U_{iso}. The influence of the anisotropic displacement parameters, as imposed on the H atoms, on the model statistics has been tested. The estimated ADPs were calculated using the IAM_U_{iso} model and the *SHADE* server (Madsen, 2006). The ADP values obtained were imposed on the H atoms in the three different charge-density models. Several cycles of refinement were carried out for these models and the corresponding IAM_U_{anis}, ELMAM_U_{anis}, ELMAM2_U_{anis} models were obtained. The results of the IAM refinement as well as the database-based models with multipole values transferred from ELMAM and ELMAM2 are presented in Table 1.

The cross-validation method based on the unbiased $R(F)_{free}$ statistical descriptor (Brünger, 1992, 1993) was used to evaluate the improvement of the database-transferred model. 5% of the reflections were selected as a test set for cross-validation of the refinements. Least-squares refinements were performed on the remaining 95% of the reflections (working set). As the free *R* factors showed large deviations, they were averaged over 20 refinements with 20 different and complementary sets of free reflections (see Table 4S).

3.9. Multipolar refinement of the AlaProAla tripeptide

The high-resolution structure of the AlaProAla tripeptide (Kalinowski *et al.*, 2007) was selected for the comparison with the new ELMAM2 databank. The final multipolar refinement as performed for the construction of the databank was selected (EXP). Two additional refinements were conducted with the charge-density parameters transferred from the ELMAM and ELMAM2 databanks. The multipolar parameters were kept fixed while the *xyz*, U_{ij} and scale-factor parameters were refined for these two models. The final statistics for all three models are listed in Table 2.

The reliability of the databanks was further assessed using the geometry from the experimental multipolar refinement. The multipolar parameters from the ELMAM, ELMAM2 and THEO models were transferred to the experimental geometry. The electrostatic interaction energy computations were conducted between dimers in the crystal of the AlaProAla tripeptide accordingly with §3.6. The correlation coefficients were calculated for the three-dimensional grids of the deformation density and the electrostatic potentials as described in §3.7. The ESP statistics were evaluated around the van der Waals surface within a shell of 2.0 Å thickness.

4. Results and discussion

In the following sections, various properties of the electron density derived from the multipolar or point charges models based on the different databanks (AMBER, ELMAM, ELMAM2) are compared with those coming from the theoretical electron-density distributions projected on the multipolar atoms model (THEO). The most important quantities describing the molecular electron density were selected in the set of peptides and aromatic molecules. Deviations from theory were tested using various statistical quantities, *e.g.* correlation coefficient, root mean square (RMS), RMS deviations (RMSD) and *R*-factor values. We decided to focus on the electron-density properties of the systems and omit the analysis of the geometric parameters, as several publications have already treated those aspects thoroughly (Jelsch *et al.*, 1998; Dittrich, Strümpel *et al.*, 2006; Dittrich *et al.*, 2008; Bąk *et al.*, 2009, 2011; Domagała *et al.*, 2011). The differences between the two multipolar databases ELMAM and ELMAM2 are expected to yield very small differences in the structure itself (Bąk *et al.*, 2011). The complete examination includes: evaluation of monopole charges, dipole and quadrupole moments, electrostatic potential and deformation electron-density comparisons on three-dimensional grids as well as electrostatic interaction energies between pairs of neighbouring molecules in the crystals. In the final sections, the refinement of the SerVal dipeptide (Moen *et al.*, 2004) and the AlaProAla tripeptide (Kalinowski *et al.*, 2007) using ELMAM- and ELMAM2-transferred models is presented to evaluate the improvement of the refinement statistics with the new database.

Table 3
RMS deviation of the atomic charges Q between the different models.

The charges are derived from the monopole valence populations: $Q = N_{val} - P_{val}$, except for the AMBER database which contains point charges.

	RMSD($Q - Q'$)			
	AMBER–THEO	ELMAM–THEO	ELMAM2–THEO	ELMAM2–ELMAM
Gly	0.422	0.311	0.107	0.210
AlaGlyAla	0.434	0.355	0.204	0.173
AlaGlyAla H_2O†	0.442	0.360	0.202	0.173
AlaProAla H_2O†	0.386	0.354	0.205	0.163
AlaTyrAla EtOH†	0.425	0.362	0.226	0.151
⟨RMSD⟩ peptides	0.422	0.348	0.189	0.174
⟨Correlation⟩	0.141	−0.020	0.454	0.822
Benzene		0.199	0.106	0.094
Catechol		0.232	0.119	0.134
Resorcinol		0.305	0.188	0.134
Dihydroxybenzoic acid		0.227	0.110	0.149
Quercetin H_2O†			0.140	
Nitrophenol alpha		0.217	0.125	0.144
Nitrophenol beta		0.217	0.127	0.144
Nitrobenzoic acid		0.233	0.110	0.171
Dinitrobenzene		0.237	0.105	0.169
Nitroaniline			0.164	
⟨RMSD⟩ aromatic		0.233	0.129	0.142
⟨Correlation⟩		0.007	0.268	0.785
⟨RMSD⟩ all	0.422	0.278	0.149	0.155
⟨Correlation⟩	0.141	−0.003	0.330	0.799

† The solvent molecule was not included in the calculations.

Table 4
RMSD values of the topological integrated AIM charges between the different charge-density models.

For the AMBER model point charges were considered.

	RMSD($Q - Q'$)			
	AMBER–THEO	ELMAM–THEO	ELMAM2–THEO	ELMAM2–ELMAM
Gly	0.361	0.122	0.041	0.106
AlaGlyAla	0.361	0.233	0.075	0.189
AlaGlyAla H_2O†	0.349	0.224	0.061	0.186
AlaProAla H_2O†	0.327	0.221	0.072	0.175
AlaTyrAla EtOH†	0.289	0.211	0.066	0.168
⟨RMSD⟩ peptides	0.337	0.202	0.063	0.165
⟨Correlation⟩	0.943	0.966	0.995	0.980
Benzene		0.229	0.056	0.173
Catechol		0.188	0.057	0.132
Resorcinol		0.187	0.068	0.126
Dihydroxybenzoic acid		0.156	0.055	0.109
Quercetin H_2O†			0.065	
Nitrophenol alpha		0.294	0.061	0.283
Nitrophenol beta		0.300	0.065	0.288
Nitrobenzoic acid		0.226	0.052	0.185
Dinitrobenzene		0.272	0.062	0.222
Nitroaniline			0.068	
⟨RMSD⟩ aromatic		0.232	0.061	0.190
⟨Correlation⟩		0.831	0.990	0.871
⟨RMSD⟩ all	0.337	0.220	0.061	0.180
⟨Correlation⟩	0.943	0.883	0.992	0.913

† The solvent molecule was not included in the calculations.

4.1. Atomic charges, dipole and quadrupole moments and AIM charges

The RMSDs of the atomic charges computed between the THEO and the database-transferred models are shown in Table 3. The monopole charge values $N_{val} - P_{val}$ were derived from the Hansen & Coppens (1978) multipole formalism. When the THEO model values are taken as reference, the largest deviations are observed for the three zwitterionic tripeptides for any databases and are the largest for the AMBER database. The lowest deviations with respect to the THEO models are observed for the ELMAM2 databank. On average, the charge deviations are about 1.8 times lower for ELMAM2 than for the former version (ELMAM) and more than 2.2 times lower than the models using the AMBER point charges. This last larger discrepancy is expected, as the same charge definition is used for all three multipolar models but a different one is used in AMBER. However, monopole charge values are not a reliable source of information about the partial charges as multipoles (notably dipoles) do also contribute to the charge transfer between atoms. The atomic charges can have different signs in the different modelling approaches. The correlation coefficient of the monopole-derived atomic charges between the databases and the theoretical model does not even reach 50%.

In order to test the agreement of the charge-density distributions in the ELMAM databases and in the theoretical model, the topological charges were calculated using the Bader (1990, 1998) AIM theory. The corresponding RMSD values are shown in Table 4. The RMSD with the THEO values is, on average, more than three times lower for ELMAM2 than for the former ELMAM database. The correlation with the THEO AIM charges is good for ELMAM and very good for ELMAM2, with correlation coefficients being equal to 0.883 and 0.992 for ELMAM and ELMAM2, respectively. It is to be noticed that the AMBER point charges correlate quite well with the THEO topological charges (r = 0.943) and poorly with the THEO $N_{val} - P_{val}$ charges. The latter charges model only part of the atomic charges, as the dipoles also transfer electron density between the atoms.

A similar situation is encountered for the dipole moments (Tables 5 and 6). The dipole moments computed for the transferred glycine and tripeptide models correlate well with the THEO model. The correlation coefficient r of the dipole magnitudes between the different models varies from 0.954 to 0.967. However, the magnitudes are, on average, much higher for AMBER and the ELMAM database. The dipole enhancement, defined as $\%\Delta\mu = 100(\mu_{Base} - \mu_{THEO})/\mu_{THEO}$, is equal to 48 and 22% for AMBER and ELMAM, respectively. The larger magnitude of the peptides' dipole moment with AMBER can be explained in part by the +1 and −1 charge imposed on N-terminus and C-terminus residues, while the same charges are smaller than unity when derived from the ELMAM (Zarychta *et al.*, 2007) and ELMAM2 databases.

In the case of ELMAM2, the average variation is small, around −2%. It is worth noting that the THEO model may also be altered because of the multipolar expansion approximation. For example, a value μ_{THEO} = 10.7 D is found for

Table 5
Dipole moments, μ (D), dipole enhancements $\Delta\mu$ (%) and deviations in directions $\angle$ (°) for the tripeptides and glycine.

The statistics between the different sets of values are given. $\%\Delta\mu = 100 \times (\mu - \mu_{THEO})/\mu_{THEO}$; $\angle$ is the angle between μ and μ_{THEO}. The discrepancy of a quantity is defined in the whole paper as: $\Delta\mu/\mu = [\sum(\mu - \mu_{THEO})^2]/[RMS(\mu)RMS(\mu_{THEO})]^{1/2}$.

	AMBER–THEO			ELMAM–THEO			ELMAM2–THEO			THEO
	μ	$\%\Delta\mu$	$\angle$	μ	$\%\Delta\mu$	$\angle$	μ	$\%\Delta\mu$	$\angle$	μ
Gly	15.0	40	5	13.4	25	1	10.6	−1	2	10.7
AlaGlyAla	32.6	31	6	27.0	9	5	21.6	−13	6	24.8
AlaGlyAla H_2O†	42.4	62	6	35.0	34	2	27.9	7	3	26.1
AlaProAla H_2O†	42.9	58	7	34.5	27	0	28.2	4	2	27.1
AlaTyrAla EtOH†	41.4	46	8	33.1	17	5	25.7	−9	4	28.3
Average		47	6		22	3		−2	3	
RMS(μ)	36			30			24			24
RMS($\mu - \mu_{THEO}$)	12			6			2			
Corr. (μ, μ_{THEO})	0.967			0.962			0.954			
$\Delta\mu/\mu$	41.6%			21.7%			8.6%			

† The solvent molecule was not included in the calculations.

Table 6
Dipole moments μ (D) and angle deviations in dipole directions $\angle$ (°) for the aromatic compounds.

	ELMAM *versus* THEO		ELMAM2 *versus* THEO		THEO
	μ	$\angle$	μ	$\angle$	μ
Catechol	5.6	20	4.0	15	2.9
Resorcinol	5.7	1	4.4	1	3.9
Dihydroxybenzoic acid	4.4	12	3.1	19	3.1
Quercetin H_2O†			4.5	9	5.9
Nitrophenol alpha	8.5	2	4.7	9	5.6
Nitrophenol beta	8.5	1	4.7	8	5.8
Nitrobenzoic acid	2.5	17	4.6	24	3.4
Nitroaniline			11.4	1	9.7
$\langle\%\Delta\mu\rangle$ enhancement	+9%		+11%		
RMS(μ)	6		6		5
RMS($\mu - \mu_{THEO}$)	2		1		
Corr. (μ, μ_{THEO})	0.845		0.890		
$\Delta\mu/\mu$	42.3%		19.8%		

† The solvent molecule was not included in the calculations.

Table 7
RMS and related statistics for quadrupole eigenvalues Q_{xx}, Q_{yy}, Q_{zz} (DÅ) of the traceless quadrupole tensor for glycine and tripeptides.

	AMBER	ELMAM	ELMAM2	THEO
RMS(Q)	55	49	37	41
RMS($Q - Q_{THEO}$)	19	11	9	
Corr. (Q, Q_{THEO})	0.966	0.988	0.979	
$\Delta Q/Q$	39.9%	24.1%	23.7%	

Table 8
RMS and related statistics for quadrupole eigenvalues Q_{xx}, Q_{yy}, Q_{zz} (DÅ) of the traceless quadrupole tensor for the aromatic molecules.

	ELMAM	ELMAM2	THEO
RMS(Q)	17	13	14
RMS($Q - Q_{THEO}$)	11	5	
Corr. (Q, Q_{THEO})	0.681	0.927	
$\Delta Q/Q$	70.1%	36.8%	

glycine (Table 5). The corresponding value in the literature was estimated to be in the range 10.15–15.76 D depending on the level of theory and the approximations used (Spackman *et al.*, 2007). The experimental value of the dipole moment for glycine, coming from high-resolution multipolar refinement, was reported to be 14.9 (3) D by Destro *et al.* (2000). Our results confirm the good correlations of the dipole moments estimated using the databases for the zwitterionic structures, which have an intrinsically large dipole moment. However, the deviations in the orientations of the dipole-moment vectors for the tripeptides (around 3–6°) are smaller than those found by Bąk *et al.* (2011) (up to 30°).

The dipole moments were also computed for the selected aromatic compounds (see Table 6). The (ELMAM, THEO) and (ELMAM2, THEO) correlation coefficients between the dipole-moment magnitudes, equal to r = 0.845 and 0.890, respectively, are lower than for the peptides. The dipole enhancement $\%\Delta\mu$ is about +42% for ELMAM and +6% for ELMAM2. The deviations in the orientations are almost three times higher than in the case of the highly polar zwitterionic moieties and reach 9 and 11° on average for ELMAM and ELMAM2, respectively. For molecules with lower dipole moments, there is a larger variability (in relative value) of the calculated dipole magnitudes and orientations with respect to the THEO model when the different models are compared.

The traceless quadrupole moment eigenvalues were calculated for the considered molecules. The corresponding statistics are shown in Table 7 and Table 8. Detailed information is presented in Table 5S and Table 6S. Comparison of the eigenvalues for the tripeptides shows the best correlation and the smallest RMSD with the THEO model for the ELMAM and the ELMAM2 databases. The AMBER model exhibits on average two times higher RMSDs. The quadrupole moments approximated by ELMAM2 are generally closer to the THEO model than from any other database, with the exception of the AlaTyrAla peptide.

The better agreement of the generalized database compared to the former one is repeated in the case of the quadrupole moments of the aromatic molecules. Comparison of the eigenvalues obtained shows that the deviations with the THEO modelling are globally twice lower for the new ELMAM2 database compared to ELMAM. The corresponding RMSD values *versus* THEO are equal to 6 and 3 DÅ for the ELMAM and ELMAM2 models, respectively. For the compounds not containing nitro groups, both correlation coefficients r are no less than 0.96.

For example, the Q_{zz} component value of benzene, as computed from ELMAM2, is equal to −10.8 DÅ, which is close both to the THEO value −9.1 DÅ and reference values

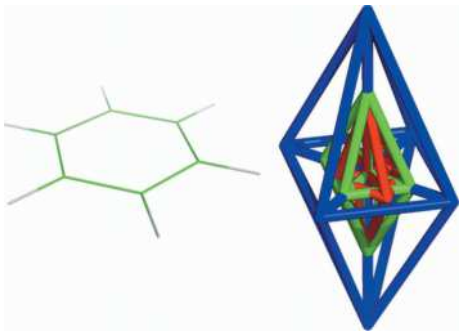

Figure 2
Principal components of the traceless quadrupole tensor of benzene are shown. In red: THEO (*CRYSTAL06* calculation); in blue: ELMAM; and in green: ELMAM2. The diagonals of the polyhedrons represent relative absolute values of the principal components of the quadrupole moment.

Table 9
Statistics for the deformation electron density $D = \Delta\rho$ with the different models.

First line of each pair: for all the peptides; second line of each pair: for the aromatic molecules. The RMS deviation and discrepancy $\Delta D/D$ with the THEO model are also given. The $\Delta\rho$ values were computed on the asymmetric unit and are in e Å^{-3}.

	ELMAM	ELMAM2	THEO
Corr. (D, $D_{\rm THEO}$)	0.893	0.969	
	0.843	0.962	
RMS($D - D_{\rm THEO}$)	0.015	0.008	
	0.018	0.009	
$\Delta D/D$	0.465	0.263	
	0.550	0.283	
RMS(D)	0.031	0.030	0.032
	0.033	0.029	0.032

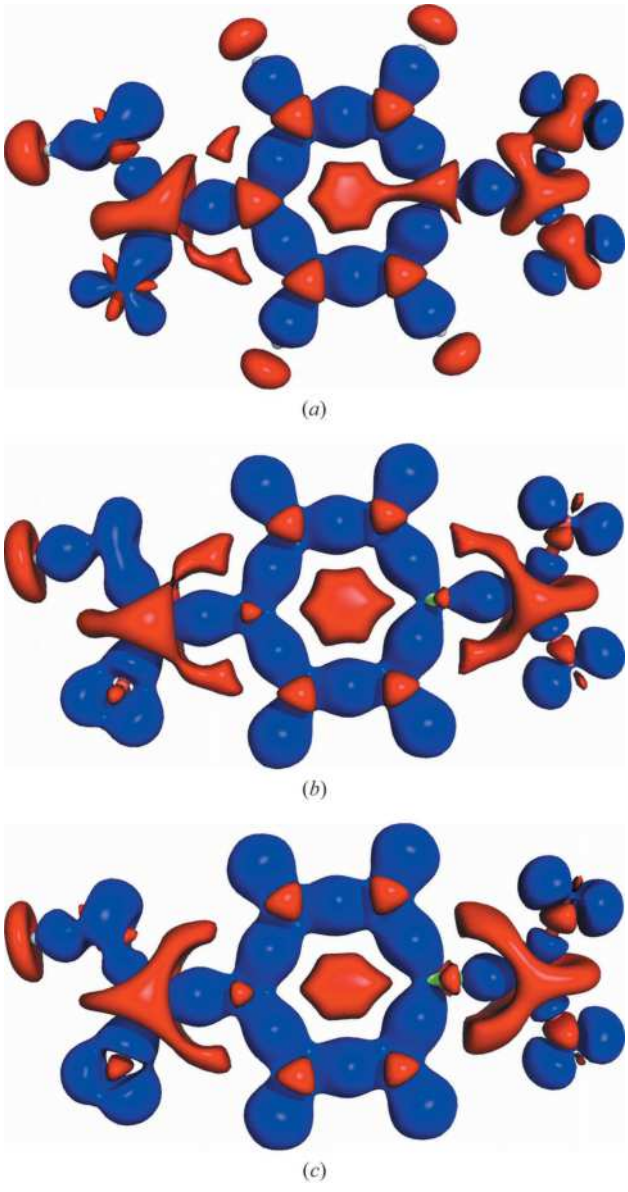

Figure 3
Deformation electron density for nitrobenzoic acid. (*a*) ELMAM, (*b*) ELMAM2 and (*c*) THEO. The isosurfaces are at +0.1 e Å^{-3} (blue) and −0.1 e Å^{-3} (red) levels.

from the literature obtained by different methods: −8.7 in gas, −8.5 in solution and between −9.7 and −12.0 from X-ray diffraction (Spackman, 1992). The Q_{zz} direction component is almost the same for the ELMAM2 and THEO models, while for ELMAM it is much larger. The directions of the *xx* and *yy* principal components of the traceless quadrupole tensor are, however, different in Fig. 2. This is due to the fact that the tensor is almost uniaxial, *i.e.* the *xx* and *yy* components for benzene are almost equal.

The relative difference of the principal components of the traceless quadrupole tensor for the benzene molecule is presented in Fig. 2. The ELMAM2 quadrupole moment of benzene shows good agreement with the THEO one, which is not the case for the ELMAM model.

When the nitro compounds are included in the set of aromatic compounds, the correlation coefficient drops significantly for ELMAM (from $r = 0.979$ to 0.445), which indicates a low quality of the electron-density parameters in the ELMAM database for the nitro group (Fig. 3). For ELMAM2, a smaller decrease is observed ($r = 0.959$ to 0.937). Including the nitro compounds increases the RMSD of the quadrupole values to 11 and 5 DÅ for ELMAM and ELMAM2, respectively. The deformation electron density of the nitro group was already found to show some discrepancy between experiment and theory (Zarychta *et al.*, 2011). The basis set employed in the *ab initio* calculations possibly may not be sufficient for accurate description of the nitro group.

4.2. Deformation electron densities

For each molecule, a three-dimensional grid of deformation electron density $\Delta\rho(\mathbf{r})$ was calculated using the electron-density distributions of the ELMAM, ELMAM2 and THEO models (see §3). The distributions of $\Delta\rho(\mathbf{r})$ were compared between the different models. The corresponding statistics for tripeptides and aromatic compounds are listed in Table 9. Detailed information is listed in Table 7S and Table 8S. Both the correlation coefficient and the RMSD values show a better accordance between ELMAM2 and THEO than between THEO and the former multipolar database. For example, the

Table 10
Comparison of the potential at the van der Waals surface between THEO and other models for the peptides.

The discrepancy $\Delta V/V$ is shown for all molecules.

	$\Delta V/V$		
	AMBER	ELMAM	ELMAM2
Glycine	0.352	0.234	0.058
AlaGlyAla	0.314	0.248	0.163
AlaGlyAla H_2O†	0.531	0.369	0.136
AlaProAla H_2O†	0.525	0.337	0.092
AlaTyrAla EtOH†	0.441	0.292	0.216
Average	0.433	0.296	0.133

† The solvent molecule was not included in the calculations.

Table 11
Comparison of the electrostatic potential at the van der Waals surface between THEO and other models for aromatic molecules.

The discrepancy $\Delta V/V$ is shown for all molecules.

	$\Delta V/V$	
	ELMAM	ELMAM2
Benzene	0.767	0.178
Catechol	0.536	0.305
Resorcinol	0.400	0.222
Dihydroxybenzoic acid	0.417	0.302
Quercetin H_2O†		0.316
Average (non-nitro)	0.530	0.265
Nitrophenol alpha	0.791	0.280
Nitrophenol beta	0.772	0.271
p-Nitrobenzoic acid	0.857	0.444
Dinitrobenzene	1.126	0.358
Nitroaniline		0.335
Average (nitro groups)	0.887	0.338
Global average	0.708	0.301

† The solvent molecule was not included in the calculations.

average RMSD values for the tripeptides are equal to 0.015 and 0.008 e Å^{-3} for ELMAM and ELMAM2, respectively.

The same situation is encountered in the case of aromatic molecules, where the RMSD values are equal to 0.018 and 0.009 e Å^{-3} for ELMAM and ELMAM2 $\Delta\rho(\mathbf{r})$ values, respectively. The correlation coefficient with $\Delta\rho(\mathbf{r})_{\rm THEO}$ is on average better (r = 0.962) for the ELMAM2 than for the ELMAM database (r = 0.843). For all the aromatic molecules tested, apart from benzene, the correlation coefficient is lower than 0.90 for the ELMAM database while for ELMAM2 it is at least 0.95. The three models display deformation electron densities of similar magnitudes; the RMS value of $\Delta\rho(\mathbf{r})_{\rm ELMAM2}$ is 5% lower compared to $\Delta\rho(\mathbf{r})_{\rm THEO}$ for both the peptidic and aromatic molecules.

For both databases, the discrepancies are higher for molecules containing nitro groups. In the case of the ELMAM models, the average discrepancy $\Delta R/R$ for molecules with and without a nitro group is equal to 0.650 and 0.460, respectively. The corresponding values for ELMAM2 are 0.335 and 0.234, respectively. A visual representation of the deformation density for the selected molecule nitrobenzoic acid is shown in Fig. 3. The figure confirms the better estimation of the deformation density in nitrobenzoic acid by the ELMAM2 databank.

4.3. Electrostatic potentials

The results of the comparison of the electrostatic potential (ESP) distributions are shown in Tables 10 and 11. The average $\Delta V/V$ values for the tripeptides indicate that the lowest deviations with respect to the THEO ESP are obtained with the ELMAM2 database. The AMBER database shows the highest deviations. Furthermore, the surface quantities characterizing the ESP on the van der Waals surface were calculated for each molecule. They are shown in Table 9S and Table 10S. Once again, the smallest differences are found between THEO and ELMAM2. The ESP at the molecular surface of the AGA (AlaGlyAla) peptide is shown in Fig. 4. The ESP shows larger values in magnitude around the negative carboxylate and the positive ammonium for the AMBER and ELMAM transfers. This is even more visible when the molecular surface is coloured by ESP differences (Fig. 5) and confirms the larger AGA dipole moment for these two models (Table 5). The nitrophenol (alpha) molecule was also selected to illustrate the difference between the databases (Fig. 6). The visual differences mirror well the quantitative information of the discrepancy factor $\Delta V/V$. The electrostatic potential distribution of nitrophenol alpha is shown for each model in Fig. 1S.

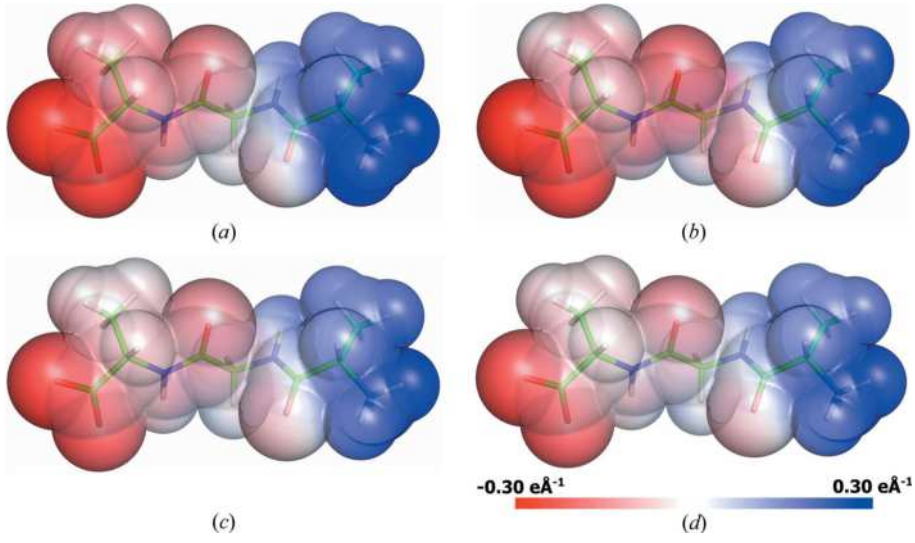

Figure 4
Electrostatic potential mapped on the van der Waals surface of the AGA molecule for (*a*) AMBER, (*b*) ELMAM, (*c*) ELMAM2 and (*d*) THEO models. The maximum negative (blue) and positive (red) values of the ESP correspond to the values −0.30 and 0.30 e Å^{-1}, respectively.

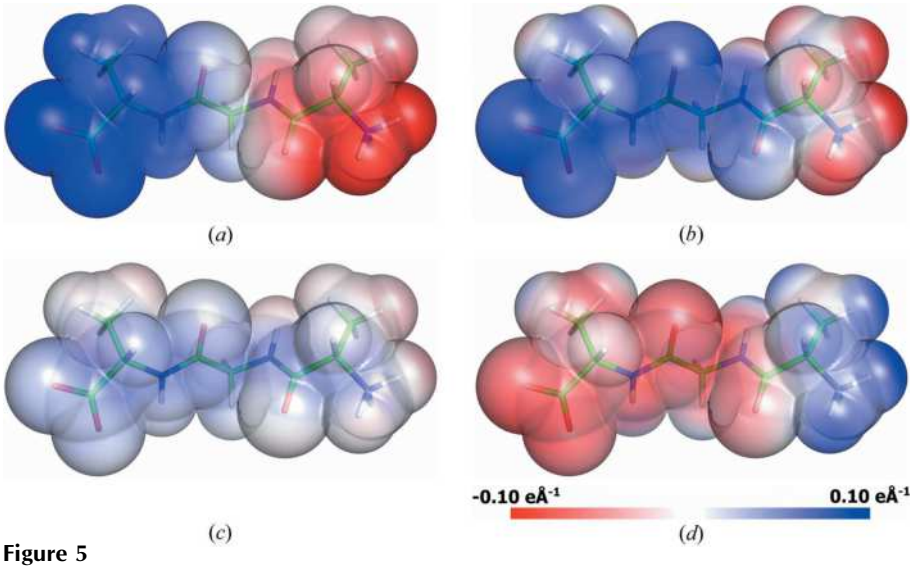

Figure 5
The differences in electrostatic potential mapped on the van der Waals surface of the AGA molecule for (*a*) THEO–AMBER, (*b*) THEO–ELMAM, (*c*) THEO–ELMAM2 and (*d*) ELMAM–ELMAM2 models. The maximum negative (blue) and positive (red) values of the ESP correspond to the values −0.10 and 0.10 e Å^{-1}, respectively. The $\Delta R/R$ values are 0.531, 0.369 and 0.136 for AMBER, ELMAM and ELMAM2, respectively.

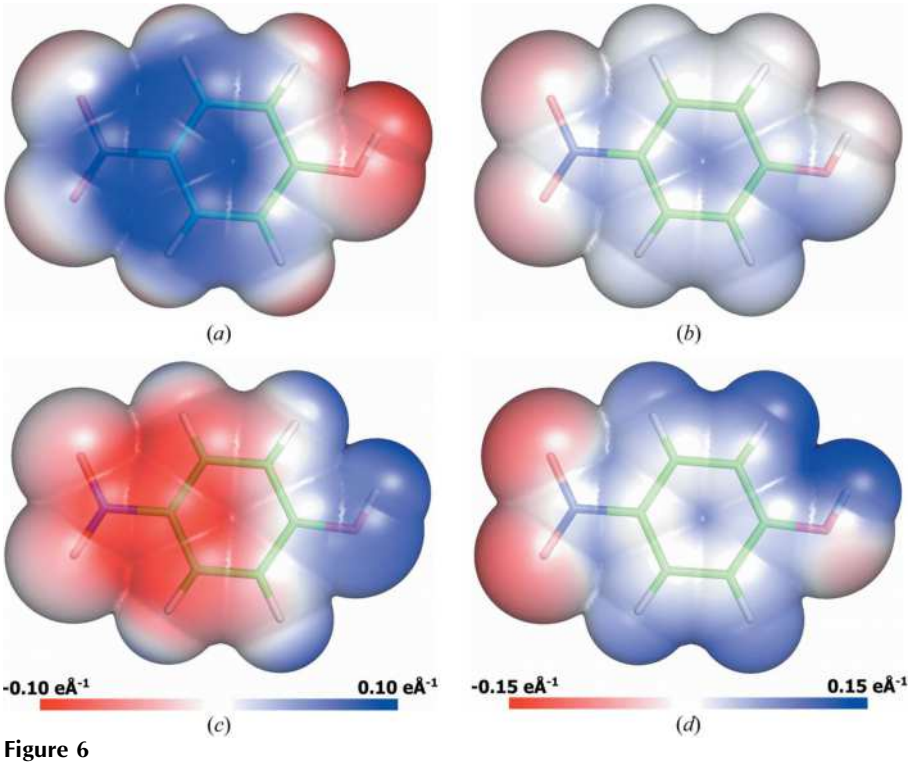

Figure 6
The differences in electrostatic potential mapped on the van der Waals surface of the nitrophenol (alpha) molecule for (*a*) THEO–ELMAM, (*b*) THEO–ELMAM2 and (*c*) ELMAM–ELMAM2 models. The maximum negative (blue) and positive (red) values of the ESP correspond to the values −0.10 and 0.10 e Å^{-1}, respectively. The total ESP of the THEO model is shown in (*d*) with a maximum ESP at ±0.15 e Å^{-1}.

4.4. Electrostatic interaction energies

The discrepancies between the electrostatic interaction energies for the tripeptides and amino acids were estimated for the three multipolar models. Calculations were carried out for the pairs of molecules in the crystals which interact by hydrogen-bond and van der Waals contacts. The solvent molecules were not included in the calculations. We selected 20 attractive electrostatic interactions between pairs of interacting molecules of glycine, AlaGlyAla, AlaProAla and AlaTyrAla. The interacting pairs of molecules with corresponding symmetry in the crystal lattice and with the values of the electrostatic interaction energies are listed in Table 12. When the THEO model is taken as reference, all the statistical descriptors (*r* correlation coefficient, RMSD and discrepancy) clearly indicate the improvement of the interaction energy estimation based on the ELMAM2 database when compared to the AMBER or ELMAM database. The AMBER point charges result in electrostatic energies which are on average 55% higher compared to the THEO model. The obtained statistics corroborate earlier results of interaction-energy calculations for 24 electrostatic interactions in dipeptides and amino acids using different multipolar databases (Bąk *et al.*, 2011). In particular, the RMSD values are equal to 95, 34 and 21 kJ mol^{-1} for AMBER, ELMAM and ELMAM2, respectively, in our study and are similar to those obtained by Bąk *et al.* (2011): 31 and 18 kJ mol^{-1} for ELMAM and ELMAM2, respectively. A slightly better correlation is observed in the present study between the two databanks. Correlation coefficients of THEO with AMBER, ELMAM and ELMAM2 are equal to 0.935, 0.973 and 0.986, respectively. Those obtained by Bąk were equal to 0.865 and 0.935 for ELMAM and ELMAM2, respectively. In both studies, theoretical electron densities obtained by the periodic calculations were projected on the same multipolar atoms model.

Similarly, the accuracy of the electrostatic interaction energies for the small aromatic compounds using the ELMAM2 database was assessed (Table 12). The energy values for the 52 attractive (according to the THEO model) interactions in the ten compounds are listed in the supplementary material. The comparison was carried out only for THEO, ELMAM and ELMAM2 models as the AMBER database does not contain

Table 12
Statistics for the attractive electrostatic interaction energies (kJ mol^{-1}) for dimers found in the crystals (contacts within the van der Waals radii of atoms).

The first line of each pair corresponds to the peptidic molecules and the second line of each pair corresponds to the aromatic molecules. The RMS values, RMS deviations, correlation coefficient and discrepancy with the THEO model are computed. The statistics for the ELMAM databank were calculated only for the transferred molecules.

	AMBER	ELMAM	ELMAM2	THEO
RMS(E)	239	179	141	154
		31	42	46
RMS($E - E_{\rm THEO}$)	95	34	21	
		22	7	
Corr. (E, $E_{\rm THEO}$)	0.935	0.973	0.985	
		0.897	0.985	
$\Delta E/E$	49.5%	20.7%	14.1%	
		56.6%	16.9%	

atom types for these kinds of molecules. The interaction energies for ELMAM were calculated only for eight structures for which appropriate atom types were available.

As for the analysis of the tripeptides, ELMAM2 exhibits better correlation with THEO than with the former database. The RMS values of the electrostatic energy are lower than in the case of peptides. The interactions are weaker in these dimers as the molecules have fewer polar groups than the peptides. The correlation and discrepancy $\Delta E/E$ values are equal to 0.985 and 17%, respectively, for the ELMAM2/THEO models. These are very close to the values obtained for the interactions between tripeptides. The accordance is lower for the ELMAM database as the corresponding values are equal to 0.897 and 57%.

4.5. Improvement of the refinement statistics

The database transfer was tested on the L-Ser–L-Val dipeptide (Moen *et al.*, 2004; §3.8). The refinement statistics for

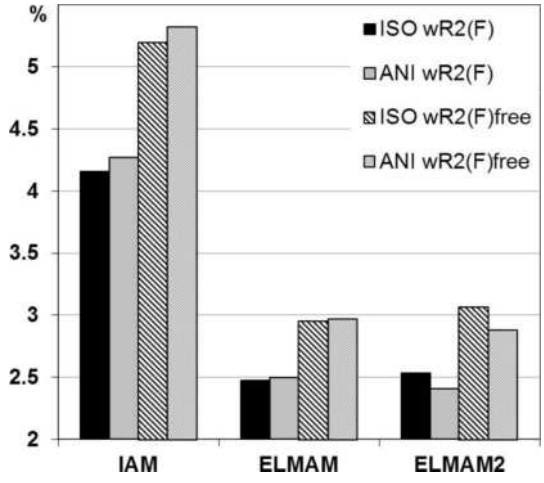

Figure 7
Cross-validation refinement statistics for the L-Ser-L-Val structure. The average $wR^2(F)$ and $wR^2(F)_{\rm free}$ values for $U_{\rm iso}$ and $U_{\rm aniso}$ models are given in %.

Table 13
Electrostatic interaction energies (kJ mol^{-1}) for the dimers of AlaProAla calculated on the experimental geometry.

	Symmetry	EXP	ELMAM	ELMAM2	THEO
AlaProAla	$\frac{1}{2}-x, 1-y, \frac{1}{2}+z$	−137	−212	−162	−163
	$-x, \frac{1}{2}+y, \frac{3}{2}-z$	−108	−140	−118	−126
	$1-x, \frac{1}{2}+y, \frac{3}{2}-z$	−119	−224	−168	−171
	$\frac{1}{2}+x, \frac{1}{2}-y, 2-z$	−17	−8	−5	−3
RMS(E)		106	169	131	134
Corr. (E, $E_{\rm EXP}$)			0.963	0.979	0.981

IAM, ELMAM and ELMAM2 clearly indicate the improvement of the overall model when the database-transferred model is used (Fig. 6): the $wR^2(F)$ value drops from 4.1 to 2.5% for both databases. This is in agreement with earlier corroborations (Bąk *et al.*, 2011; Domagała *et al.*, 2011; Zarychta *et al.*, 2011). The values of the maximum and minimum residual electron density are reduced by a factor of almost 2 when the database models are used.

Fig. 7 shows the cross-validation free $wR^2(F)$ statistical descriptor (Brünger, 1992, 1993). As expected, the free R factor drops significantly when the IAM refinement is augmented with the parameters transferred from the database. The $wR^2(F)_{\rm free}$ value drops from 5.2% to a value around 3.0%.

When an isotropic thermal motion of the H atoms is considered, the ELMAM_$U_{\rm iso}$ model is slightly better than ELMAM2_$U_{\rm iso}$ in terms of statistical descriptors. This might be due to the fact that the first and second version of the database were constructed by using, respectively, isotropic and anisotropic descriptions of the H-atom thermal motion. Including full anisotropic displacement parameters from the *SHADE* server (Madsen, 2006) reverses the situation. Now the ELMAM2_$U_{\rm anis}$ becomes the best model, which is in accordance with the anisotropic modelling of H atoms in the construction of the ELMAM2 database. ELMAM2 does also contain q_{3z^2-1} quadrupole components for the H atoms, in contrast to ELMAM, and results therefore in a better separation of the thermal motion of atoms and electron density in ELMAM2. Globally, the best crystallographic R and free R factors are obtained with an anisotropic description of the thermal motion of the H atoms and using the ELMAM2 database.

4.6. Databases *versus* AlaProAla experimental charge density

The refinement statistics for ELMAM and ELMAM2 indicate that, of course, the full multipolar refinement (EXP) of the tripeptide AlaProAla is the best approach (Table 2). However, the reliability of the ELMAM2 model is much higher compared to ELMAM as indicated by the lower R factor (more than 0.2% improvement). This is consistent with the previous example of the L-Ser–L-Val refinement, as here the anisotropic description of the H-atom thermal motion was also applied. The correlation of the electrostatic interaction energy values with respect to the EXP model is improved from ELMAM to ELMAM2 but the THEO models yield the best

Table 14
Statistics for the deformation density $\Delta\rho$ and electrostatic potential V of the AlaProAla peptide for different models.

The underlined values on the diagonal of the table are the RMS values of the deformation densities (RMS$\Delta\rho$), computed in the asymmetric unit (e Å^{-3}). The values of the correlation coefficients between the deformation electron densities $\Delta\rho$ are given in italics below the diagonal. The values of the correlation coefficients between the electrostatic potentials (V) are given above the diagonal. The statistics were computed on the grids of the deformation density and the electrostatic potential using experimental geometry.

	ELMAM	ELMAM2	EXP	THEO
ELMAM	0.031	0.993	0.960	0.989
ELMAM2	*0.915*	0.029	0.958	0.997
EXP	*0.900*	*0.962*	0.034	0.960
THEO	*0.880*	*0.966*	*0.937*	0.032

agreement. Particular interactions appear to be weaker in the EXP model. The RMS(E) value is much lower for the experimental model (Table 13). The correlations of the ESP values with the EXP model are similar for all the models and close to 0.960 (Table 14). The best correlations between values of the deformation densities $\Delta\rho$ are observed between the ELMAM2 and EXP models. The $\Delta\rho$ correlation between ELMAM and the experimental model is the lowest. Globally, for all properties, the lowest deviations from the experimental model are found for the ELMAM2 or THEO models, whereas the ELMAM transfer shows lower agreement.

5. Concluding remarks

A new database of experimental multipolar atom types has been constructed – ELMAM2 – which is an improved version of the former ELMAM databank. The aim of the expansion was to describe also the electron-density distributions for common organic molecules. For the construction of ELMAM2, an automated matching strategy and assignment of atom types is applied, based on the chemical environment of the tested atoms. The optimal local axes were used to minimize the number of multipole values stored in the databank. The ELMAM2, ELMAM and AMBER databases were compared using a sample of 15 crystal structures that contained both peptides and small aromatic molecules. The following properties, related to the electron-density distributions, have been extensively compared: monopole and AIM charges, dipole and quadrupole moments, deformation density distributions, electrostatic potentials and electrostatic interaction energies.

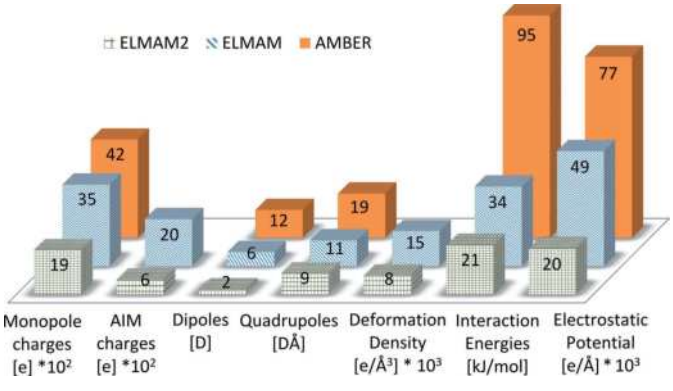

Figure 8
Global RMS deviations from the THEO model presented for AMBER (orange, solid colour), ELMAM (blue, diagonal stripes) and ELMAM2 (green, crosshatching). The average values were computed for the tripeptides and glycine, but not for the organic molecules.

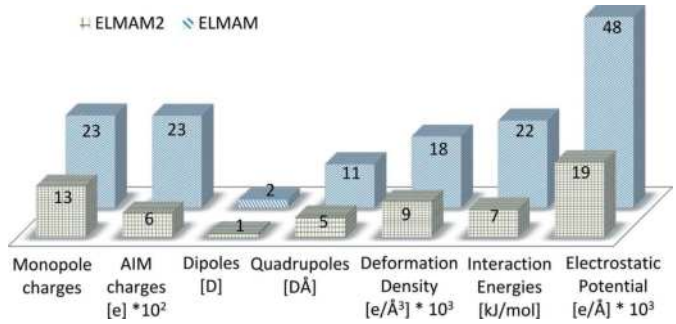

Figure 9
Global RMS deviations between the THEO model and ELMAM (blue, diagonal stripes) and ELMAM2 (green, crosshatching) transfers. The average values were computed for the aromatic molecules.

The results clearly show greater accuracy of the new ELMAM2 database in comparison to ELMAM or the point charges AMBER databank. The good predictability of the new database was evaluated using the root mean square deviations, discrepancy factors $\Delta R/R$ and correlation descriptors with respect to the theoretical multipolar model. The RMS deviations to the THEO values show on average a reduction close to 50% for the ELMAM2 model transferred to the peptides in comparison to ELMAM and more than 68% in comparison to the AMBER model (summing all the properties) (see Fig. 8). The same situation is repeated for the aromatic organic molecules. More than 57% of reduction in RMS deviations are observed for ELMAM2 with respect to ELMAM models (Fig. 9). The multipolar model was extended to the quadrupole level (q_{3z^2-1}) for H atoms; this modelling, similar to the theoretical one, is probably one of the reasons why there are smaller differences between the THEO and ELMAM2 results.

References

Ahmed, M., Noureen, S., Gros, P. C., Guillot, B. & Jelsch, C. (2011). *Acta Cryst.* C**67**, o329–o333.
Allen, F. H., Kennard, O., Watson, D. G., Brammer, L., Orpen, A. G. & Taylor, R. (1987). *J. Chem. Soc. Perkin Trans. 2*, S1–S19.
Allen, F. H., Watson, D. G., Brammer, L., Orpen, A. G. & Taylor, R. (2006). *International Tables for Crystallography*, Vol. C, ch. 9.5, pp. 790–811. Dordrecht: Kluwer Academic Publishers.
Bacon, G. E., Curry, N. A. & Wilson, S. A. (1964). *Proc. R. Soc. London Ser. A*, **279**, 98–110.
Bacon, G. E. & Jude, R. J. (1973). *Z. Kristallogr.* **138**, 19–40.
Bader, R. F. W. (1990). *Atoms in Molecules: a Quantum Theory*. Oxford University Press.
Bader, R. F. W. (1998). *J. Phys. Chem. A*, **102**, 7314–7323.

Bąk, J. M., Domagała, S., Hübschle, C., Jelsch, C., Dittrich, B. & Dominiak, P. M. (2011). *Acta Cryst.* A**67**, 141–153.
Bąk, J. M., Dominiak, P. M., Wilson, C. C. & Woźniak, K. (2009). *Acta Cryst.* A**65**, 490–500.
Becke, A. D. (1988). *J. Chem. Phys.* **88**, 2547–2553.
Becke, A. D. (1993). *J. Chem. Phys.* **98**, 5648–5652.
Benabicha, F., Pichon-Pesme, V., Jelsch, C., Lecomte, C. & Khmou, A. (2000). *Acta Cryst.* B**56**, 155–165.
Birkedal, H., Madsen, D., Mathiesen, R. H., Knudsen, K., Weber, H.-P., Pattison, P. & Schwarzenbach, D. (2004). *Acta Cryst.* A**60**, 371–381.
Bouhmaida, N., Bonhomme, F., Guillot, B., Jelsch, C. & Ghermani, N. E. (2009). *Acta Cryst.* B**65**, 363–374.
Brock, C. P., Dunitz, J. D. & Hirshfeld, F. L. (1991). *Acta Cryst.* B**47**, 789–797.
Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
Brünger, A. T. (1993). *Acta Cryst.* D**49**, 24–36.
Case, D. A. *et al.* (2008). *AMBER 10*. University of California, San Francisco, USA.
Chęcińska, L., Förster, D., Morgenroth, W. & Luger, P. (2006). *Acta Cryst.* C**62**, o454–o457.
Chen, Y.-S., Stash, A. I. & Pinkerton, A. A. (2007). *Acta Cryst.* B**63**, 309–318.
Cohen, D. E., Benedict, J. B., Morlan, B., Chiu, D. T. & Kahr, B. (2007). *Crystal Growth Des.* **7**, 492–495.
Coppens, P. (1997). *X-ray Charge Densities and Chemical Bonding*. New York: Oxford University Press.
Coppens, P., Abramov, Y., Carducci, M., Korjov, B., Novozhilova, I., Alhambra, C. & Pressprich, M. R. (1999). *J. Am. Chem. Soc.* **121**, 2585–2593.
Dahaoui, S., Jelsch, C., Howard, J. A. K. & Lecomte, C. (1999). *Acta Cryst.* B**55**, 226–230.
Destro, R., Marsh, R. E. & Bianchi, R. (1988). *J. Phys. Chem.* **92**, 966–973.
Destro, R., Roversi, P., Barzaghi, M. & Marsh, R. E. (2000). *J. Phys. Chem. A*, **104**, 1047–1054.
Dittrich, B., Hübschle, C. B., Holstein, J. J. & Fabbiani, F. P. A. (2009). *J. Appl. Cryst.* **42**, 1110–1121.
Dittrich, B., Hübschle, C. B., Luger, P. & Spackman, M. A. (2006). *Acta Cryst.* D**62**, 1325–1335.
Dittrich, B., Hübschle, C. B., Messerschmidt, M., Kalinowski, R., Girnt, D. & Luger, P. (2005). *Acta Cryst.* A**61**, 314–320.
Dittrich, B., Koritsánszky, T., Grosche, M., Scherer, W., Flaig, R., Wagner, A., Krane, H. G., Kessler, H., Riemer, C., Schreurs, A. M. M. & Luger, P. (2002). *Acta Cryst.* B**58**, 721–727.
Dittrich, B., Koritsánszky, T. & Luger, P. (2004). *Angew. Chem. Int. Ed.* **43**, 2718–2721.
Dittrich, B., McKinnon, J. J. & Warren, J. E. (2008). *Acta Cryst.* B**64**, 750–759.
Dittrich, B., Munshi, P. & Spackman, M. A. (2007). *Acta Cryst.* B**63**, 505–509.
Dittrich, B. & Spackman, M. A. (2007). *Acta Cryst.* A**63**, 426–436.
Dittrich, B., Strumpel, M., Schäfer, M., Spackman, M. A. & Koritsánszky, T. (2006). *Acta Cryst.* A**62**, 217–223.
Dittrich, B., Weber, M., Kalinowski, R., Grabowsky, S., Hübschle, C. B. & Luger, P. (2009). *Acta Cryst.* B**65**, 749–756.
Domagała, S. & Jelsch, C. (2008). *J. Appl. Cryst.* **41**, 1140–1149.
Domagała, S., Korybut-Daszkiewicz, B., Straver, L. & Woźniak, K. (2009). *Inorg. Chem.* **48**, 4010–4020.
Domagała, S., Munshi, P., Ahmed, M., Guillot, B. & Jelsch, C. (2011). *Acta Cryst.* B**67**, 63–78.
Dominiak, P. M., Grech, E., Barr, G., Teat, S., Mallinson, P. & Woźniak, K. (2003). *Chem. Eur. J.* **9**, 963–970.
Dominiak, P. M., Volkov, A., Dominiak, A. P., Jarzembska, K. N. & Coppens, P. (2009). *Acta Cryst.* D**65**, 485–499.
Dominiak, P. M., Volkov, A., Li, X., Messerschmidt, M. & Coppens, P. (2007). *J. Chem. Theory Comput.* **3**, 232–247.
Dovesi, R., Saunders, V. R., Roetti, C., Orlandol, R., Zicovich-Wilson, C. M., Pascale, F., Civalleri, B., Doll, K., Harrison, N. M., Bush, I. J., D'Arco, Ph. & Llunell, M. (2008). *CRYSTAL06 1.0*. Version 1_0_2. University of Turin, Italy.
Espinosa, E., Lecomte, C., Molins, E., Veintemillas, S., Cousson, A. & Paulus, W. (1996). *Acta Cryst.* B**52**, 519–534.
Faerman, C. H. & Price, S. L. (1990). *J. Am. Chem. Soc.* **112**, 4915–4926.
Förster, D., Messerschmidt, M. & Luger, P. (2005). *Acta Cryst.* C**61**, o420–o421.
Fournier, B., Bendeif, el-E., Guillot, B., Podjarny, A., Lecomte, C. & Jelsch, C. (2009). *J. Am. Chem. Soc.* **131**, 10929–10941.
Fronczek, F. R., Kim, K. K. & Strongin, R. M. (2002). Private communication (refcode: CATCOL13).
Ghermani, N. E., Spasojević-de Biré, A., Bouhmaida, N., Ouharzoune, S., Bouligand, J., Layre, A., Gref, R. & Couvreur, P. (2004). *Pharm. Res.* **21**, 598–607.
Groth, P. (1980). *Acta Chem. Scand. A*, **334**, 229–230.
Guillot, B., Jelsch, C., Podjarny, A. & Lecomte, C. (2008). *Acta Cryst.* D**64**, 567–588.
Guillot, B., Muzet, N., Artacho, E., Lecomte, C. & Jelsch, C. (2003). *J. Phys. Chem. B*, **107**, 9109–9121.
Guillot, B., Viry, L., Guillot, R., Lecomte, C. & Jelsch, C. (2001). *J. Appl. Cryst.* **34**, 214–223.
Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* A**34**, 909–921.
Hariharan, P. C. & Pople, J. A. (1973). *Theor. Chim. Acta*, **28**, 213–222.
Hohenberg, P. & Kohn, W. (1964). *Phys. Rev. B*, **136**, 864–871.
Hoser, A. A., Dominiak, P. M. & Woźniak, K. (2009). *Acta Cryst.* A**65**, 300–311.
Howard, J. A. K., Mahon, M. F., Raithby, P. R. & Sparkes, H. A. (2009). *Acta Cryst.* B**65**, 230–237.
Hübschle, C. B., Dittrich, B., Grabowsky, S., Messerschmidt, M. & Luger, P. (2008). *Acta Cryst.* B**64**, 363–374.
Jelsch, C., Guillot, B., Lagoutte, A. & Lecomte, C. (2005). *J. Appl. Cryst.* **38**, 38–54.
Jelsch, C., Pichon-Pesme, V., Lecomte, C. & Aubry, A. (1998). *Acta Cryst.* D**54**, 1306–1318.
Kalinowski, R., Dittrich, B., Hübschle, C. B., Paulmann, C. & Luger, P. (2007). *Acta Cryst.* B**63**, 753–767.
Koritsanszky, T. S. & Coppens, P. (2001). *Chem. Rev.* **101**, 1583–1627.
Kulkarni, G. U., Kumaradhas, P. & Rao, C. N. R. (1998). *Chem. Mater.* **10**, 3498–3505.
Le Page, Y. & Gabe, E. J. (1979). *J. Appl. Cryst.* **12**, 464–466.
Lebedev, V. I. & Laikov, D. N. (1999). *Dokl. Math.* **59**, 477–481.
Lee, C., Yang, W. & Parr, R. G. (1988). *Phys. Rev. B*, **37**, 785–789.
Luger, P., Messerschmidt, M., Scheins, S. & Wagner, A. (2004). *Acta Cryst.* A**60**, 390–396.
Lutz, M., Spek, A. L., van der Geer, E. P. L., van Koten, G. & Klein Gebbink, R. J. M. (2008). *Acta Cryst.* C**64**, o87–o90.
Madsen, A. Ø. (2006). *J. Appl. Cryst.* **39**, 757–758.
Madsen, A. Ø., Sørensen, H. O., Flensburg, C., Stewart, R. F. & Larsen, S. (2004). *Acta Cryst.* A**60**, 550–561.
Martin, A. & Pinkerton, A. A. (1998). *Acta Cryst.* B**54**, 471–477.
Meents, A., Dittrich, B., Johnas, S. K. J., Thome, V. & Weckert, E. F. (2008). *Acta Cryst.* B**64**, 42–49.
Moen, A., Frøseth, M., Görbitz, C. H. & Dalhus, B. (2004). *Acta Cryst.* C**60**, o564–o565.
Munshi, P. & Guru Row, T. N. (2002). *Acta Cryst.* B**58**, 1011–1017.
Munshi, P. & Guru Row, T. N. (2005*a*). *Crystallogr. Rev.* **11**, 199–241.
Munshi, P. & Guru Row, T. N. (2005*b*). *J. Phys. Chem. A*, **109**, 659–672.
Munshi, P. & Guru Row, T. N. (2006*a*). *Crystal Growth Des.* **6**, 708–718.
Munshi, P. & Guru Row, T. N. (2006*b*). *Acta Cryst.* B**62**, 612–626.
Munshi, P., Thakur, T. S., Guru Row, T. N. & Desiraju, G. R. (2006). *Acta Cryst.* B**62**, 118–127.
Murray, J. S., Peralta-Inga, Z. & Politzer, P. (2000). *Int. J. Quantum Chem.* **80**, 1216–1223.

Murray, J. S. & Politzer, P. (1998). *J. Mol. Struct. (Theochem)*, **425**, 107–114.
Muzet, N., Guillot, B., Jelsch, C., Howard, E. & Lecomte, C. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 8742–8747.
Nieger, M. (2007). Private communication (refcode: NANILI23).
Ogawa, K., Noda, Y., Lüthi, T. & Buergi, H.-B. (2006). Personal communication.
Overgaard, J. & Hibbs, D. E. (2004). *Acta Cryst.* A**60**, 480–487.
Padiyar, G. S. & Seshadri, T. P. (1996). *Acta Cryst.* C**52**, 1693–1695.
Parrish, D., Zhurova, E. A., Kirschbaum, K. & Pinkerton, A. A. (2006). *J. Phys. Chem. B*, **110**, 26442–26447.
Pichon-Pesme, V., Lachekar, H., Souhassou, M. & Lecomte, C. (2000). *Acta Cryst.* B**56**, 728–737.
Pichon-Pesme, V., Lecomte, C. & Lachekar, H. (1995). *J. Phys. Chem.* **99**, 6242–6250.
Rees, B. (1976). *Acta Cryst.* A**32**, 483–488.
Rodrigues, B. L., Tellgren, R. & Fernandes, N. G. (2001). *Acta Cryst.* B**57**, 353–358.
Scheins, S., Dittrich, B., Messerschmidt, M., Paulmann, C. & Luger, P. (2004). *Acta Cryst.* B**60**, 184–190.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Slouf, M., Holy, A., Petříček, V. & Cisarova, I. (2002). *Acta Cryst.* B**58**, 519–529.
Sørensen, H. O., Stewart, R. F., McIntyre, G. J. & Larsen, S. (2003). *Acta Cryst.* A**59**, 540–550.
Spackman, M. A. (1992). *Chem. Rev.* **92**, 1769–1797.
Spackman, M. A. (1997). *Annu. Rep. Prog. Chem. Sect. C Phys. Chem.* **94**, 177–207.
Spackman, M. A., Munshi, P. & Dittrich, B. (2007). *ChemPhysChem*, **8**, 2051–2063.
Sparkes, H. A., Brayshaw, S. K., Weller, A. S. & Howard, J. A. K. (2008). *Acta Cryst.* B**64**, 550–557.
Tang, W., Sanville, E. & Henkelman, G. (2009). *J. Phys. Comput. Mater.* **21**, 084204–084210.
Tonogaki, M., Kawata, T., Ohba, S., Iwata, Y. & Shibuya, I. (1993). *Acta Cryst.* B**49**, 1031–1039.
Treutler, O. & Ahlrichs, R. (1995). *J. Chem. Phys.* **102**, 346–354.
Tsirelson, V. G. & Ozerov, R. P. (1996). *Electron Density and Bonding in Crystals.* Bristol: Institute of Physics Publishing.
Volkov, A., Abramov, Y. A. & Coppens, P. (2001). *Acta Cryst.* A**57**, 272–282.
Volkov, A., Abramov, Y., Coppens, P. & Gatti, C. (2000). *Acta Cryst.* A**56**, 332–339.
Volkov, A., Li, X., Koritsánszky, T. S. & Coppens, P. (2004). *J. Phys. Chem. A*, **108**, 4283–4300.
Volkov, A., Messerschmidt, M. & Coppens, P. (2007). *Acta Cryst.* D**63**, 160–170.
Wiest, R., Pichon-Pesme, V., Bénard, M. & Lecomte, C. (1994). *J. Phys. Chem.* **98**, 1351–1362.
Zarychta, B., Pichon-Pesme, V., Guillot, B., Lecomte, C. & Jelsch, C. (2007). *Acta Cryst.* A**63**, 108–125.
Zarychta, B., Zaleski, J., Kyzioł, J., Daszkiewicz, Z. & Jelsch, C. (2011). *Acta Cryst.* B**67**, 250–262.
Zhurov, V. V., Zhurova, E. A., Chen, Y.-S. & Pinkerton, A. A. (2005). *J. Appl. Cryst.* **38**, 827–829.
Zhurova, E. A., Martin, A. & Pinkerton, A. A. (2002). *J. Am. Chem. Soc.* **124**, 8741–8750.
Zhurova, E. A., Matta, C. F., Wu, N., Zhurov, V. V. & Pinkerton, A. A. (2006). *J. Am. Chem. Soc.* **128**, 8849–8861.
Zhurova, E. A. & Pinkerton, A. A. (2001). *Acta Cryst.* B**57**, 359–365.