Statistical analysis of multipole-model-derived structural parameters and charge-density properties from high-resolution X-ray diffraction experiments
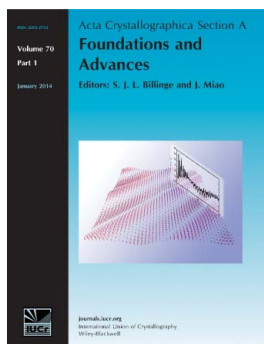
Radosław Kamiński, Sławomir Domagała, Katarzyna N. Jarzembska, Anna A. Hoser, W. Fabiola Sanjuan-Szklarz, Matthias J. Gutmann, Anna Makal, Maura Malińska, Joanna M. Bąk and Krzysztof Woźniak

Acta Crystallographica Section A: Foundations of Crystallography covers theoretical and fundamental aspects of the structure of matter. The journal is the prime forum for research in diffraction physics and the theory of crystallographic structure determination by diffraction methods using X-rays, neutrons and electrons. The structures include periodic and aperiodic crystals, and non-periodic disordered materials, and the corresponding Bragg, satellite and diffuse scattering, thermal motion and symmetry aspects. Spatial resolutions range from the subatomic domain in charge-density studies to nanodimensional imperfections such as dislocations and twin walls. The chemistry encompasses metals, alloys, and inorganic, organic and biological materials. Structure prediction and properties such as the theory of phase transformations are also covered.

Crystallography Journals Online is available from **journals.iucr.org**

research papers

Statistical analysis of multipole-model-derived structural parameters and charge-density properties from high-resolution X-ray diffraction experiments

Radosław Kamiński,[a]*‡§ Sławomir Domagała,[a]*§ Katarzyna N. Jarzembska,[a] Anna A. Hoser,[a] W. Fabiola Sanjuan-Szklarz,[a] Matthias J. Gutmann,[b] Anna Makal,[a] Maura Malińska,[a] Joanna M. Bąk[a] and Krzysztof Woźniak[a]

[a]Department of Chemistry, University of Warsaw, Pasteura 1, 02-093 Warszawa, Poland, and [b]ISIS Neutron and Muon Source, Science and Technology Facilities Council, Rutherford Appleton Laboratory, Harwell Oxford, Didcot, Oxfordshire OX11 0QX, UK. Correspondence e-mail: rkaminski@chem.uw.edu.pl, slawdom@chem.uw.edu.pl

A comprehensive analysis of various properties derived from multiple high-resolution X-ray diffraction experiments is reported. A total of 13 charge-density-quality data sets of α-oxalic acid dihydrate ($C_2H_2O_4 \cdot 2H_2O$) were subject to Hansen–Coppens-based modelling of electron density. The obtained parameters and properties were then statistically analysed yielding a clear picture of their variability across the different measurements. Additionally, a computational approach (*CRYSTAL* and *PIXEL* programs) was utilized to support and examine the experimental findings. The aim of the study was to show the real accuracy and interpretation limits of the charge-density-derived data. An investigation of raw intensities showed that most of the reflections (60–70%) fulfil the normality test and the lowest ratio is observed for weak reflections. It appeared that unit-cell parameters are determined to the order of 10^{-3} Å (for cell edges) and 10^{-2} ° (for angles), and compare well with the older studies of the same compound and with the new 100 K neutron diffraction data set. Fit discrepancy factors are determined within a 0.5% range, while the residual density extrema are about ± 0.16 (3) e Å^{-3}. The geometry is very well reproducible between different data sets. Regarding the multipole model, the largest errors are present on the valence shell charge-transfer parameters. In addition, symmetry restrictions of multipolar parameters, with respect to local coordinate systems, are well preserved. Standard deviations for electron density are lowest at bond critical points, being especially small for the hydrogen-bonded contacts. The same is true for kinetic and potential energy densities. This is also the case for the electrostatic potential distribution, which is statistically most significant in the hydrogen-bonded regions. Standard deviations for the integrated atomic charges are equal to about 0.1 e. Dipole moments for the water molecule are comparable with the ones presented in various earlier studies. The electrostatic energies should be treated rather qualitatively. However, they are quite well correlated with the *PIXEL* results.

1. Introduction

The experimental charge-density field has undoubtedly reached its mature stage (Coppens, 2005). Having crystalline samples of an appropriate quality and modern X-ray equipment, it is possible to collect very good high-resolution diffraction data sets. Such data are indispensable to quantitatively model electron-density distribution within crystals.

Nowadays, because of the advanced development of instruments, techniques and software, charge-density studies of simple organic molecules have become almost routine, similar to structural analysis. Nevertheless, there are still a number of challenging scientific problems in the field, such as charge-density distribution modelling of crystal structures containing heavy atoms (Zhurov, Zhurova & Pinkerton, 2011; Zhurov, Zhurova, Stash *et al.*, 2011*b*; Schmøkel *et al.*, 2013), crystals exhibiting a high degree of ionicity (*e.g.* minerals) (Schmøkel *et al.*, 2012), treatment of disorder (Bąk *et al.*, 2009; Meindl *et al.*, 2009), description of inner atomic electronic shells (Fischer

et al., 2011), proper treatment of hydrogen atoms (Hoser *et al.*, 2009; Munshi *et al.*, 2008) and thermal motion deconvolution (Zhurov, Zhurova, Stash *et al.*, 2011*a*), among others. More complex crystal cases may exhibit a variety of interesting and useful properties (mechanical, optical *etc.*) and thus are worth exploring. Extraction of reliable electron-density information for such samples often causes severe difficulties. Consequently, new approaches are constantly being developed, as for example the most recent X-ray constrained wavefunction fitting technique (Jayatilaka & Grimwood, 2001). There are also various ideas for improving the most widely used Hansen–Coppens model (Hansen & Coppens, 1978). Modelling procedures should, however, be designed with care, as they could also become a source of significant bias affecting the final results (Hoser *et al.*, 2009; Spackman & Byrom, 1997*a*; El Haouzi *et al.*, 1996; Howard *et al.*, 1995). Therefore, we do need to be aware of the *limitations* of the applied tools and methods in the widest context possible.

Recently, a number of studies explored the deficiencies of the multipole model (Bąk *et al.*, 2009, 2011, 2012; Poulain-Paul *et al.*, 2012; Dittrich *et al.*, 2012). Nevertheless, it should be noted that recorded diffracted-beam intensities also constitute a source of experimental errors, being further propagated across all stages of electron-density elucidation. Thus, to obtain an overall view of the method's real shortcomings all of these factors should be taken into account. This could be done by applying statistical methods to an experimental sample. Quite recently, Grabowsky *et al.* (2009) have investigated statistical aspects of the transferability principles regarding charge-density distribution properties. Additionally, the authors briefly compared their results with the literature examples of standard deviations obtained from several (from two to four) charge-density measurements of the same compound. However, the mentioned analyses are not fully reliable because of the few data sets used in the study. Hence, in this paper, we indicate the precision of the determined structural and charge-density parameters on the basis of multiple high-resolution data collections of oxalic acid. The main idea of our study can be summarized by the following question: 'What are the differences between the data sets, geometrical parameters and resulting charge-density models of the same chemical system derived from different high-resolution X-ray measurements which we could classify as acceptable?'

We have collected a total of 13 charge-density-quality data sets for α-oxalic acid dihydrate ($C_2H_2O_4 \cdot 2H_2O$) crystals. This constitutes a good statistical sample, especially considering the very time-consuming experiments. There are several reasons for the choice of oxalic acid. Firstly, the crystals of this compound can be easily grown and, importantly, with good reproducibility of their quality. The other reason is that oxalic acid is probably the best charge-density-investigated compound so far. Since the very beginning, the crystals of this compound were used as a reference. The earliest data collections were conducted by Stevens *et al.* (1979), followed by Coppens *et al.* (1984) who carried out an excellent survey of both X-ray and neutron data sets collected across the globe

(four X-ray and five neutron data sets are reported therein). The most up-to-date charge-density study of the kind was performed by Martin & Pinkerton (1998). This was also the first application of the CCD detector for experimental electron-density elucidation. Our investigation is supplemented by brief comparison with these older results. This allows us to gain a deeper insight into the experimental reproducibility and various factors that influence the charge-density X-ray data quality.

2. Methods

2.1. Crystallization

The studied oxalic acid dihydrate was purchased from the L.P.P-H 'OCH' Lublin chemical company as a crystalline powder. The compound quickly dissolves in water and, upon evaporation at room temperature under a microscope, forms crystals suitable for X-ray examination. Crystals of the size appropriate for a neutron diffraction experiment were grown by a very slow evaporation of water at room temperature from a medium-size beaker.

2.2. Neutron diffraction

A single-crystal neutron diffraction experiment was performed at the SXD beamline (Keen *et al.*, 2006), at the neutron spallation source in the ISIS facility (Oxfordshire, UK). For the purpose of the experiment a single crystal of dimensions $4 \times 2 \times 1$ mm was attached to an Al pin with an adhesive Al tape and cooled to 100 K in a closed-cycle refrigerator with He as an exchange gas. A total of six crystal orientations were recorded with an exposure time of 1 h per orientation. Data reduction was accomplished using the locally available *SXD* program (Keen *et al.*, 2006). The crystal structure (with anisotropic hydrogen-atom thermal parameters) was then refined in the *JANA* package (Petříček *et al.*, 2006) using the time-of-flight data. It is also worth noting that the severe extinction was successfully modelled by the Becker–Coppens approach (Becker & Coppens, 1974*a,b*, 1975). A summary of refinement information follows: total No. of reflections: 2630; No. of reflections with $I \geq 3\sigma(I)$: 1583; index ranges: $-21 \leq h \leq 17$, $-5 \leq k \leq 8$, $-30 \leq l \leq 40$; No. of parameters: 65; discrepancy factors: $R[F]$ $[I \geq 3\sigma(I)]$ = 6.71%, $wR[F^2]$ $[I \geq 3\sigma(I)]$ = 6.45%, $R[F]$ (all data) = 10.83%, $wR[F^2]$ (all data) = 6.83%, goodness of fit (GoF) $[I \geq 3\sigma(I)]$ = 1.64; nuclear density largest extrema: $-4.75/+4.40$ fm Å^{-3}; extinction parameter (Gaussian isotropic): 2.36 (6). The CIF file is available in the supplementary material,[1] or can be retrieved from the Cambridge Structural Database (Allen, 2002) (deposition number: CCDC 930687). Hereafter the model is abbreviated as sxd1.

2.3. X-ray data collection

High-resolution single-crystal X-ray diffraction experiments were performed on three different experimental setups.

---

[1] Supplementary material for this paper are available from the IUCr electronic archives (Reference: PC5031).

Table 1
Parameters characterizing the X-ray data collection and multipolar refinement for all data sets.

Standard deviations of unit-cell volume are estimated from least-squares procedures. In the diffractometer description Bruker(m), Bruker(o) and KUMA denote Bruker AXS machines with monochromator or multilayer optics, and our upgraded KUMA diffraction (now Agilent Technologies) setups, respectively.

Model	oxa1	oxa2	oxa3	oxa4	oxa5	oxa6	oxa7
Diffractometer	Bruker(o)	Bruker(o)	Bruker(o)	Bruker(o)	Bruker(o)	Bruker(o)	Bruker(o)
Volume, V (Å^3)	244.63 (2)	244.54 (1)	244.46 (2)	243.79 (1)	244.92 (2)	244.39 (2)	244.54 (3)
d_{calc} (g cm^{-3})	1.712	1.712	1.713	1.718	1.710	1.714	1.713
Completeness (%)	95	98	99	91	97	97	95
$(\sin\theta/\lambda)_{max}$ (Å^{-1})	1.18	1.17	1.15	1.15	1.17	1.17	1.15
θ range (°)	3.44–57.00	3.45–55.96	3.45–54.82	3.45–54.77	3.445–56.053	3.45–55.93	3.54–54.81
Absorption coefficient, μ (mm^{-1})	0.181	0.181	0.181	0.182	0.181	0.181	0.181
Index ranges	$-13 \leq h \leq 14$ $-7 \leq k \leq 8$ $-20 \leq l \leq 27$	$-14 \leq h \leq 13$ $-7 \leq k \leq 7$ $-27 \leq l \leq 27$	$-13 \leq h \leq 14$ $-7 \leq k \leq 8$ $-26 \leq l \leq 27$	$-13 \leq h \leq 13$ $-6 \leq k \leq 7$ $-26 \leq l \leq 24$	$-14 \leq h \leq 14$ $-8 \leq k \leq 7$ $-25 \leq l \leq 27$	$-13 \leq h \leq 13$ $-7 \leq k \leq 8$ $-26 \leq l \leq 27$	$-13 \leq h \leq 13$ $-7 \leq k \leq 7$ $-26 \leq l \leq 18$
No. of reflections collected/unique	17234/3202	27377/3167	24573/3051	12130/2791	23539/3169	24801/3154	11034/2939
R_{int}	1.97%	4.87%	2.60%	2.15%	3.83%	2.58%	2.24%
No. of reflections with $I \geq 3\sigma(I)$†	2690 (84%)	1998 (63%)	2588 (85%)	2205 (79%)	2765 (87%)	2720 (86%)	2499 (85%)
No. of parameters/ restraints	177/3	177/3	177/3	177/3	177/3	177/3	177/3
$R[F]$ $[I \geq 3\sigma(I)]$	1.415	1.990	1.386	1.504	1.651	1.283	1.363
$R[F^2]$ $[I \geq 3\sigma(I)]$	2.232	2.037	2.294	2.079	3.426	3.089	2.704
$wR[F]$ $[I \geq 3\sigma(I)]$	2.094	2.165	2.278	1.895	2.379	1.835	1.852
$wR[F^2]$ $[I \geq 3\sigma(I)]$	3.931	4.154	4.039	3.598	4.370	3.414	3.572
GoF $[I \geq 3\sigma(I)]$	1.319	1.061	1.450	1.065	1.467	1.222	1.206
$\varrho_{res}^{min/max}$ (e Å^{-3})	−0.13/+0.13	−0.19/+0.17	−0.15/+0.13	−0.13/+0.14	−0.20/+0.20	−0.22/+0.20	−0.16/+0.14

Model	oxa8	oxa9	oxa10	oxa11	oxa12	oxa13
Diffractometer	Bruker(m)	Bruker(m)	Bruker(m)	Bruker(o)	Bruker(o)	KUMA
Volume, V (Å^3)	244.58 (1)	245.00 (1)	245.09 (2)	244.67 (2)	244.86 (2)	245.33 (1)
d_{calc} (g cm^{-3})	1.712	1.709	1.709	1.712	1.710	1.707
Completeness (%)	89	96	85	96	99	> 99
$(\sin\theta/\lambda)_{max}$ (Å^{-1})	1.20	1.20	1.00	1.03	1.17	1.13
θ range (°)	3.45–58.52	3.44–58.52	3.44–45.48	3.45–47.19	3.45–56.42	3.44–53.26
Absorption coefficient, μ (mm^{-1})	0.181	0.181	0.181	0.181	0.181	0.181
Index ranges	$-14 \leq h \leq 13$ $-7 \leq k \leq 8$ $-22 \leq l \leq 28$	$-14 \leq h \leq 14$ $-8 \leq k \leq 8$ $-27 \leq l \leq 26$	$-10 \leq h \leq 10$ $-7 \leq k \leq 3$ $-17 \leq l \leq 23$	$-12 \leq h \leq 12$ $-7 \leq k \leq 7$ $-24 \leq l \leq 24$	$-14 \leq h \leq 14$ $-8 \leq k \leq 8$ $-27 \leq l \leq 27$	$-13 \leq h \leq 13$ $-7 \leq k \leq 7$ $-26 \leq l \leq 26$
No. of reflections collected/unique	11581/3149	15900/3422	7347/1759	24619/2248	21907/3266	22748/2929
R_{int}	2.32%	4.61%	2.58%	1.78%	2.41%	2.11%
No. of reflections with $I \geq 3\sigma(I)$†	2282 (72%)	2755 (81%)	1434 (82%)	2125 (95%)	2892 (89%)	2394 (82%)
No. of parameters/ restraints	177/3	177/3	177/3	177/3	177/3	177/3
$R[F]$ $[I \geq 3\sigma(I)]$	1.656	1.924	1.602	1.034	1.226	1.378
$R[F^2]$ $[I \geq 3\sigma(I)]$	2.713	3.560	2.453	2.556	2.050	3.032
$wR[F]$ $[I \geq 3\sigma(I)]$	1.835	2.252	1.841	1.848	1.916	1.592
$wR[F^2]$ $[I \geq 3\sigma(I)]$	3.681	4.512	3.718	3.417	3.514	3.207
GoF $[I \geq 3\sigma(I)]$	0.941	1.090	1.000	1.465	1.252	0.988
$\varrho_{res}^{min/max}$ (e Å^{-3})	−0.18/+0.14	−0.21/+0.20	−0.17/+0.13	−0.17/+0.13	−0.11/+0.12	−0.18/+0.14

† Values given in parentheses indicate the percentage of all unique reflections.

Details are given in Table 1. In each case a new crystal was grown and mounted on the goniometer head, installed on the four-circle goniometer, and cooled to a temperature of 100 K (temperature instability is an additional factor affecting derived charge-density parameters). Depending on the crystal overall quality and its orientation, the data sets were collected using different strategies optimized for each case (only ω scans; scan width 0.5°; exposure time from 3 to 10 s on low angles, 45 to 70 s on high angles). Data sets were integrated with the respective diffractometer software: *APEX2* (Bruker AXS Inc., 2013) or *CRYSALIS* (Agilent Technologies, 2010). Further data reduction, correction application and merging were done with the *SORTAV* program (Blessing, 1987, 1989, 1995a).

2.4. Multipole refinements

All multipole refinements, within the Hansen–Coppens formalism, were performed in the *MOPRO* suite (Jelsch *et al.*, 2005) combined with the new version of the University at

Buffalo Data Bank (UBDB) (Jarzembska & Dominiak, 2012) and neutron diffraction data. All deformation parameters are defined with respect to their local Cartesian coordinate systems, which were automatically assigned by the *LSDB* program (Jarzembska & Dominiak, 2012; Volkov *et al.*, 2004). All refinements were based on F and only the reflections fulfilling the $I \geq 3\sigma(I)$ condition were taken into account (see below). In all cases statistical weights were used (*i.e.*, for the ith reflection $w_i = 1/\sigma_i^2$). Initial atomic coordinates (x, y, z) and anisotropic displacement parameters (U_{ij}'s) for each atom were taken from the spherical refinement stage, whereas initial multipolar and contraction–expansion parameters were transferred from the UBDB with the aid of the *LSDB* program. The charges of both moieties (*i.e.*, acid and water molecules) were initially set neutral. The multipole expansion was truncated at the hexadecapole ($l_{max} = 4$) and quadrupole ($l_{max} = 2$) levels for all non-hydrogen and hydrogen atoms, respectively. Each atom was assigned with core and spherical-valence scattering factors derived from Su and Coppens atomic wavefunctions (Su & Coppens, 1998). The refinement of each oxalic acid model was performed in two stages, (*a*) and (*b*), which are summarized below.

(*a*) During the first stage of the refinement procedure the *SHADE*-server-derived temperature factors for hydrogen atoms were used (Madsen, 2006). The oxalic acid and water molecules were entered into the *SHADE* server separately, so as to estimate the atomic displacement parameters (ADPs) using solely the riding approximation approach [a reliable TLS fit (Schomaker & Trueblood, 1968; Cruickshank, 1956; Sands, 1982) is impossible for such small molecules]. At this initial refinement stage the local symmetry restrictions were strictly followed. Such a procedure was found to be *essential* as it prevented falling into the false minimum during the least-square refinement procedure [the false minimum problem is discussed *e.g.* by Watkin (2008) or in the *JANA* program 'cookbook' (Petříček *et al.*, 2006)]. $X-$H bond lengths ($X =$ non-hydrogen atom) were standardized to neutron-normalized distances taken from the literature and restrained using $\sigma = 0.001$. This approach has recently been successfully tested on a variety of uracil derivatives, providing geometrical results comparable with the corresponding theoretical periodic computation and neutron studies, and seems to be the best choice in the absence of neutron diffraction data (Hoser *et al.*, 2009; Jarzembska *et al.*, 2012). Additionally, hydrogen-atom U_{iso} parameters (*i.e.*, isotropic thermal parameters) were restrained to the value of $1.5 \times U_{eq}^{O}$ using $\sigma = 0.01$ (where an appropriate restraint weight is equal to $1/\sigma^2$). The refinement was carried out according to the following scheme: (i) scale factor (which was also refined in all other stages); (ii) atomic coordinates; (iii) atomic coordinates and ADPs; (iv) a few cycles of *SHADE* estimation and point (iii) until convergence is reached; (v) multipole parameters in an incremental manner; (vi) all population, positional and thermal parameters simultaneously; (vii) a few cycles of *SHADE* estimation and point (vi); (viii) 50 cycles of block refinement of (vi) and κ parameters for non-hydrogen atoms; (ix) all parameters

simultaneously; (x) a few cycles of *SHADE* estimation and point (ix).

(*b*) In the second stage of the refinement procedure the neutron diffraction data were included. $X-$H bond lengths are restrained to the values derived from the sxd1 model. The hydrogen-atom ADPs from neutron studies are scaled using the Blessing method (Blessing, 1995*b*) and subjected as fixed to the multipolar refinement performed as follows: (xi) refinement of all population, positional and thermal parameters simultaneously; (xii) 50 cycles of block refinement of point (xi) and κ parameters for non-hydrogen atoms; (xiii) all possible parameters simultaneously. Beginning from point (xi) the local symmetry is not preserved and all populations are allowed to vary.

The final parameters characterizing refinements for all data sets are summarized in Table 1. CIF files for each refinement are present in the supplementary material, or can be retrieved from the Cambridge Structural Database (deposition numbers: CCDC 925544–925554, CCDC 925558, CCDC 925559). All *identical* refinements were performed smoothly and reached an acceptable convergence of maximal shift over the estimated standard deviation equal to around 10^{-3} or less (except for the oxa7 model, where the value of 10^{-2} was obtained, though all parameters are reasonable). This ensures that all parameters derived from multipole models are comparable one to another in the sense of statistical analysis of various factors treated as a sample. In other words, the differences between models are not 'hidden' by the deficiencies of the multipole model itself. It is also worth mentioning that in all cases the Hirshfeld rigid-bond test (Hirshfeld, 1976) is fulfilled, ensuring a proper deconvolution of electron density and thermal motion features.

Finally, the use of the $I/\sigma(I)$ cutoff during multipole refinement needs a brief discussion. This issue was described some time ago by Watkin (2008) in his excellent review paper; therefore here we shall focus solely on the cutoff effect related directly to the subject of the current contribution. Our definition of 'weak data' is that the $I/\sigma(I)$ ratio is below 3.0. Such a cutoff criterion is based on statistical considerations related to the normal distribution (the normality of raw intensity data is discussed later on in this article). However, in general, inclusion or omission of weak data in the refinement may bias the results. In fact, the correct procedure whether to reject or include such reflections should depend on their reciprocal vector length. If reflections are located in the low-angle region they should be included as they are related to the atomic valence populations. Rejection of such reflections obviously may bias to some extent the result of the charge-density refinement. When a reflection belongs to the high-angle shell, in our opinion, it is a matter of choice. Such a reflection shall mostly affect the positional and thermal motion parameters. Following the findings of Hirshfeld & Rabinovich (1973) and Seiler *et al.* (1984) regarding the data cutoff, it can be concluded that in general differences on the refined parameters should be very small. Further considerations on a resolution-dependent reflection rejection were presented by Schwarzenbach *et al.* (1989).

research papers

Keeping in mind all the above factors, we collected our data sets in a careful way so as to obtain a minimum number of weak low-angle reflections. In fact they do not exceed 20% (out of several hundreds of recorded reflections). This number is further reduced during the outlier rejection stage in *SORTAV*. To ensure that our refinement results are not substantially biased by the applied cutoff, we performed all charge-density refinements both with and without application of this condition. The results of the statistical analysis of multipole populations are presented in the supplementary material. In general, there are no statistically significant differences (being within 1σ margin) in terms of the multipole populations between the two sets of results.

2.5. Derived properties

The experimental charge-density distributions obtained by the multipolar approach were analysed by means of Bader's Quantum Theory of Atoms in Molecules (QTAIM) (Bader, 1994). The *VMOPRO* module (part of the *MOPRO* package) was used for Fourier syntheses and bond critical point (BCP) evaluation. The same program was used for Fourier map visualization and computation of dipole and quadrupole moments. Visualization of bond paths and surfaces was accomplished with the *MOPROVIEWER* program (Guillot, 2012). Neighbouring molecules, used for the bond path search, were built up using the *CLUSTERGEN* program (Kamiński *et al.*, 2013). Electric dipoles and quadrupoles were derived from the multipole populations using the formulas given by Spackman (1992) and Coppens (1997). 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. Evaluation and integration of atomic basins were performed with the *WINXPRO* program (Stash & Tsirelson, 2002, 2005).

The experimental electrostatic interaction energy values ($E_{\text{elst}}^{\text{exp}}$) between various bi-molecular complexes composed of oxalic acid or water molecules were computed with the *VMOPRO* module as an integration over the whole space of the electron density of molecule A multiplied by the electrostatic potential generated by molecule B (or reciprocally):

$$E_{\text{elst}}^{\text{exp}} = \int \varrho_A(\mathbf{r})\varphi_B(\mathbf{r})\,\mathrm{d}\mathbf{r} = \int \varrho_B(\mathbf{r})\varphi_A(\mathbf{r})\,\mathrm{d}\mathbf{r}.$$

Such integrals were computed using a numerical integration method based on a spherical grid around selected atoms. The Gauss–Chebyshev (Becke, 1988*a*) and Lebedev & Laikov quadratures (Lebedev & Laikov, 1999) 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 smaller than, or equal to, the sum of their van der Waals radii.

2.6. *PIXEL* calculations

Lattice and bi-molecular motif energy computation was performed using the *OPIX* program (Gavezzotti, 2002, 2003). All the 13 charge-density-quality structures of oxalic acid were used to calculate the corresponding molecular electron-density grids by standard quantum chemical methods using *GAUSSIAN* (Frisch *et al.*, 2009) at the MP2/6-31G** level of theory (Møller & Plesset, 1934; Krishnan *et al.*, 1980). The electron-density model of each molecule was analysed in the *OPIX* package. It was possible to extract the interaction energy values, and their breakdown into electrostatic, polarization, dispersion and repulsion components, for bi-molecular motifs present in the crystal structure. The electrostatic and polarization energy terms were compared directly with the results obtained on the basis of the multipolar model.

2.7. Periodic computations

Cohesive energy computations were performed with the *CRYSTAL* program package (Dovesi *et al.*, 2005, 2009) at the DFT(B3LYP) level of theory (Lee *et al.*, 1988; Becke, 1988*b*). Two types of molecular all-electron basis sets were used, *i.e.* either 6-31G** (Krishnan *et al.*, 1980) or pVTZ (Dunning, 1989), which is specified in the text. Both the Grimme dispersion correction (Grimme, 2004, 2006) and a correction for basis set superposition error (BSSE) were applied (Civalleri *et al.*, 2008). Ghost atoms were selected up to 5 Å distance from the considered molecule in a crystal lattice and were used for the BSSE estimation. The evaluation of Coulomb and exchange series was controlled by five thresholds, set arbitrarily to the values of 10^{-7}, 10^{-7}, 10^{-7}, 10^{-7}, 10^{-25}. The shrinking factor was equal to 8, which refers to 170 **k**-points in the irreducible Brillouin zone and assures the full convergence of the total energy. The cohesive energy (E_{coh}) was calculated as described below (Civalleri *et al.*, 2008):

$$E_{\text{coh}} = \frac{1}{Z}E_{\text{bulk}} - E_{\text{mol}},$$

where E_{bulk} is the total energy of a system (calculated per unit cell), E_{mol} is the energy of a molecule extracted from the bulk and Z stands for the number of molecules in the unit cell. For comparative purposes two geometry optimizations at the DFT(B3LYP) level of theory with the pVTZ and 6-31G** basis sets were carried out. During the optimization procedure atomic coordinates were varied while cell parameters were kept fixed.

3. Statistical methods background

3.1. Basic statistical descriptors

In this section the most common and useful statistical descriptors are summarized (Brandt, 1998). In the case of a set of measurements $\{x_i\}$ ($i = 1, \ldots, n$) a *sample mean* is defined as

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i,$$

Table 2
Numerical values of unit-cell parameters for all oxalic acid structures (specific values are given with the estimated standard deviation computed by the diffractometer software).

The corresponding values obtained from the neutron diffraction experiment are given for comparison and are not included in the statistical analysis. Shorthand notation for the statistically averaged X-ray data values is given in bold.

Model	Unit-cell parameters (Å, °)			
	a	b	c	β
oxa1	6.0931 (6)	3.4921 (2)	11.8409 (7)	103.842 (2)
oxa2	6.0893 (2)	3.4965 (1)	11.8311 (4)	103.884 (2)
oxa3	6.0910 (3)	3.4928 (2)	11.8372 (6)	103.901 (2)
oxa4	6.0889 (2)	3.4870 (1)	11.8299 (3)	103.930 (1)
oxa5	6.0939 (3)	3.4965 (2)	11.8441 (5)	103.951 (2)
oxa6	6.0909 (3)	3.4916 (2)	11.8401 (6)	103.935 (2)
oxa7	6.0950 (6)	3.4881 (3)	11.8462 (9)	103.841 (3)
oxa8	6.0903 (2)	3.4959 (1)	11.8348 (3)	103.917 (1)
oxa9	6.0954 (2)	3.4948 (1)	11.8499 (3)	103.9382 (9)
oxa10	6.0988 (2)	3.4954 (1)	11.8455 (4)	103.928 (2)
oxa11	6.0920 (3)	3.4919 (2)	11.8481 (5)	103.898 (2)
oxa12	6.0938 (4)	3.4929 (2)	11.8533 (7)	103.948 (2)
oxa13	6.0989 (1)	3.49746 (6)	11.8494 (2)	103.921 (2)
Mean values	$\bar{a}$	$\bar{b}$	$\bar{c}$	$\bar{\beta}$
	6.093 (3)	**3.493 (3)**	**11.842 (7)**	**103.91 (4)**
sxd1	6.118 (2)	3.504 (2)	11.883 (4)	103.93 (3)
sxd2	6.100 (2)	3.494 (2)	11.849 (4)	103.90 (3)

which stands for the estimator of the expectation value. The *sample standard deviation* is expressed as

$$s_x = \left[\frac{1}{n-1} \sum_{i=1}^{n} (\bar{x} - x_i)^2 \right]^{1/2} .$$

It is worth noting here that, whereas s_x^2 constitutes the *unbiased* estimator of the sample variance, s_x itself in general cannot be treated as the unbiased estimator of the standard deviation. Under the assumption that the sample obeys the normal distribution it is, though, possible to obtain the unbiased estimator by multiplying s_x by a certain correction factor [which can be derived by the use of Cochran's theorem (Cochran, 1934)]. However, this approach was not utilized here as the correction factor is very close to unity ($\simeq 0.979$ for 13 data sets) so its application does not affect the final conclusions.

3.2. Normality test

Numerous advanced statistical methods concern the so-called normal distribution, which means that a given statistical population can be described by a Gaussian-type function. However, to apply the appropriate statistical tests, one needs to check whether the examined statistical sample can be treated as normally distributed. The well known Shapiro–Wilk test (Shapiro & Wilk, 1965) is the best method for such verification. In this approach one may reject the null hypothesis saying that a given property comes from a normally distributed population, or state that there is not enough information to reject it. The test algorithm is as follows: (i) sort the values in an increasing order; (ii) compute the test statistic as

$$W = \frac{1}{s_x^2 (n-1)} \left[\sum_{k=1}^{K} a_{k,n} (x_{n-k+1} - x_k) \right]^2 ,$$

where K equals $n/2$ or $(n-1)/2$ (for even- or odd-populated samples, respectively), whereas the $a_{k,n}$ values are tabulated in standard statistical tables; (iii) the null hypothesis is rejected when $W \leq W(\alpha, n)$, where the critical values $W(\alpha, n)$ are taken from statistical tables (where α is a given significance level). In our case of 13 data sets, the critical values equal 0.814, 0.866 and 0.889 for α equal to 1, 5 and 10%, respectively.

4. Results and discussion

In the beginning of the analysis and discussion it should be noted that we were not able to perform a fully systematic study of various factors (crystal, X-ray radiation source, detector *etc.*) affecting the final set of obtained charge-density quantities. All 13 data sets were measured within several years mainly for the purpose of diffractometer alignment for charge-density studies. Therefore, we limited ourselves to investigating the overall variability of different experimentally derived parameters. Nevertheless, the study can be extended in the future when more data sets, preferably measured using other X-ray sources, are available.

4.1. Unit-cell parameters

The above-described statistical methods were first applied to analyse, compare and contrast a set of unit-cell parameters determined from different measurements. The corresponding numerical values are given in Table 2. It is worth noting here that the unit-cell setting used in this study ($P2_1/c$) is different to that used in previous works ($P2_1/n$). The unit-cell setting was changed, following Dunitz (1996) and Feast *et al.* (2009), to minimize parameter correlations in oblique-angle crystal systems.

It is a well known fact that the precision of CCD detectors is much worse than that obtained with point detectors (Herbstein, 2000). Our results for estimated standard deviations are, though, comparable with the early estimates given in the literature (Paciorek *et al.*, 1999). They are even larger than the ones reported by Guzei *et al.* (2008) for 2-dimethylsulfur-anylidene-1,3-indanedione, where the sample standard deviation value (averaged over 223 data sets) amounts to about 0.0005 Å. It is also clearly seen that the calculated sample standard deviations are at least an order of magnitude larger than the values resulting from the least-square fits of orientation matrices. In addition, they are rather large, especially in the case of the β unit-cell angle.

It is noticeable that the obtained standard deviation values are quite similar to the corresponding parameters resulting from the orientation matrix least-squares fit to the neutron diffraction data. It should be stressed that neutron data are processed quite differently when compared to the X-ray diffraction case. The sxd1 model unit-cell parameters presented in Table 2 are computed as averaged values over all detector settings using the centre-of-mass position of peak

centroids on the detector pixels and along time-of-flight as determined by the initial peak searching routine. The lattice parameters denoted as sxd2 in Table 2 use the fitted position of the peaks as determined during the integration. They are very close to the sample means computed from the X-ray diffraction experiments. We found out that for the neutron data the influence of the unit-cell parameters (sxd1 or sxd2) on various structural parameters (*e.g.* hydrogen-atom ADPs) is rather negligible. These results indicate a very good agreement of the two methods, keeping in mind that the neutron diffraction technique is also significantly affected by factors such as extinction, absorption, neutron-beam wavelength distribution or flux inaccuracies.

In order to analyse the nature of the experimental errors, the Shapiro–Wilk test was applied. This yielded the conclusion that for all of the unit-cell parameters the normal distribution hypothesis cannot be rejected at any of the considered significance levels. It is then justifiable to treat the errors as a result of random fluctuations of environmental conditions during each experiment (*e.g.* slightly unstable temperature, goniometer angle inaccuracy, detector distance *etc.*). It is worth stressing that the unit-cell constants used in the study were obtained for precise and accurate charge-density-quality measurements. The high-angle data allow a very precise description of the atomic positions and thus also of the unit-cell parameters. Consequently, in the case of many standard low-resolution structural data, the experimental standard deviations may be even larger. Importantly, the so-called unconstrained unit-cell refinements (*i.e.*, with the α and γ angles not being fixed at $90°$) led to values of right angles in the monoclinic system not significantly different from $90°$. Finally, comparison of the above results with the ones presented by Coppens *et al.* (1984) and Martin & Pinkerton (1998) (see Table 2 in the original paper) shows a very good agreement. The unit-cell parameters in the previous works deviate by about 4σ at worst (for the *a* parameter) from our values. This confirms a good reproducibility of unit-cell parameters, even recorded with the older experimental setups.

On the whole, in the case of a single X-ray diffraction experiment we suggest using the least-squares estimates of unit-cell parameters, even though they are clearly lower than the real ones [this conclusion matches the statement made by Farrugia & Evans (2005) and Guzei *et al.* (2008)]. We believe it is better to show the results that are at least mathematically sound than the ones that are based on rather vague reasoning. In addition, our results for unit-cell parameters have serious implications for various analyses of structural phase transitions or time-resolved processes. In such cases particular attention has to be paid to unit-cell changes.

4.2. Raw intensity data analysis

Careful analysis of the raw intensity data is a desirable procedure in every charge-density study (Kuntzinger *et al.*, 1999). For the purpose of this study, we decided to provide a statistical analysis of the collected raw data. Thus, in order to apply the statistical methods to raw intensities we first brought

them to the common (absolute) scale by applying the respective scale factors from the final multipole refinements. Subsequently, we extracted reflection indices being common for all 13 data sets. This yielded us a set of 1662 reflections with resolution up to about 1.0 Å^{-1} (Table 1).

At first, we analysed raw intensities (*I*) and their estimated standard deviations [$\sigma(I)$] across all data sets. Typical scatter plots are shown in Fig. 1, where the second data set is compared to the 12th one (all other comparisons are very much the same). For all data sets the *I versus I* relation is almost exactly linear (with the linear determination coefficient, R^2, being in the range from 0.918 to 0.999). In the case of standard deviations discrepancies from linearity are more pronounced ($R^2 = 0.374$–0.847); however, these are less representative as they are only *estimated* during the integration procedures.

The most important issue is the normality of such distributions. To answer this question we utilized the Shapiro–Wilk test based on the raw intensities. The resultant plots are shown
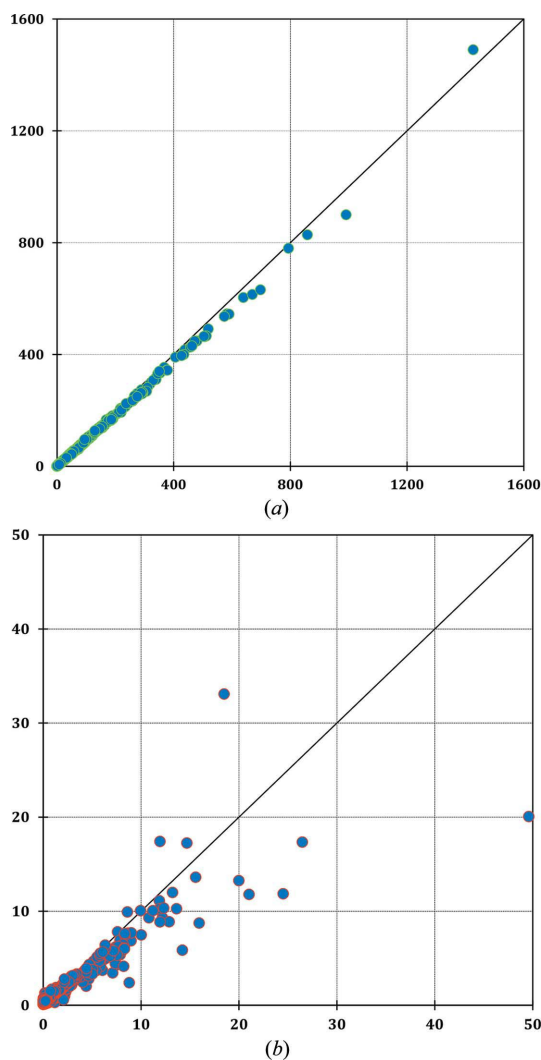

Figure 1
Example plots of (a) I_{set2} *versus* I_{set12} ($R^2 = 0.997$) and (b) $\sigma(I_{\text{set2}})$ *versus* $\sigma(I_{\text{set12}})$ ($R^2 = 0.831$) relationships for the raw (scaled) data sets. The black solid line is the reference line.

in Fig. 2. A plot of the intensity-based Shapiro–Wilk test with respect to the resolution shows that at least more than 50% of reflections pass the normality test. The lowest acceptance ratio is observed for the most high-resolution reflections, and for the rest this ratio varies from 60 to 70%. Interesting features are also present in the plot illustrating the results of the mentioned test with respect to the intensity values. Clearly, a maximum is observed. The largest deviations from normality are observed for the very weak reflections (25% pass the test), then the ratio of acceptance increases to 80%, but then it drops down to about 62% for the most intense reflections. The low acceptance ratio for the very weak reflections can be understood as resulting from the increased measurement problems. In the case of strong reflections the background estimation and detector characteristics could play an important role.

4.3. Fit quality and residual density

The very first quantities derived for all the charge-density models are the R factors and residual density distributions.
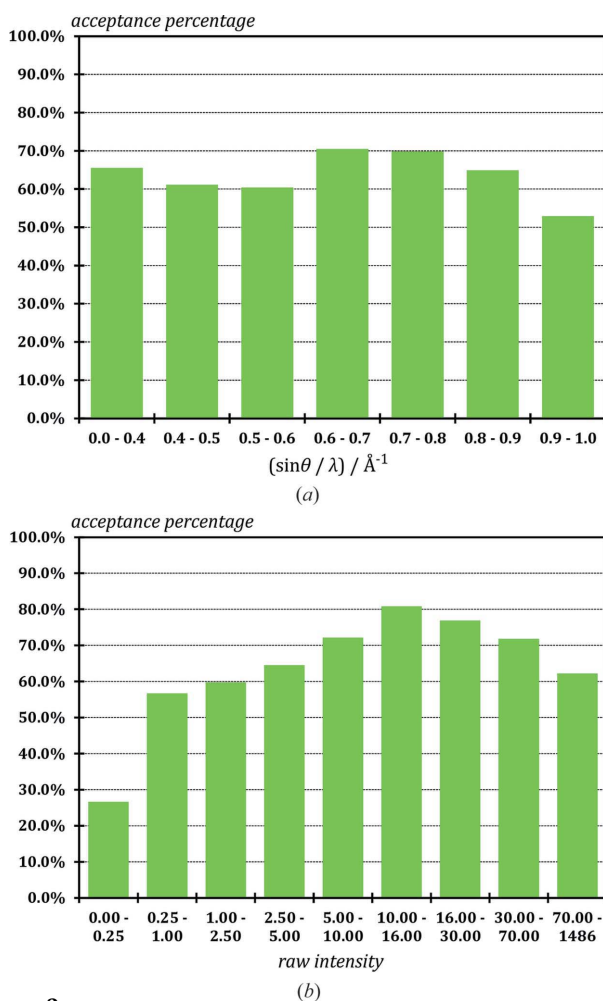

Figure 2
Histograms showing the acceptance ratios ($\alpha = 10\%$) for the Shapiro–Wilk test on the reflection raw intensity values *versus* (*a*) resolution and (*b*) raw (scaled and averaged) intensity.

Table 3
Mean values and estimated standard deviations for all computed R factors and GoF values from Table 1.

Descriptor	Value
$R[F]$ $[I \geq 3\sigma(I)]$	1.5 (3)%
$R[F^2]$ $[I \geq 3\sigma(I)]$	2.6 (5)%
$wR[F]$ $[I \geq 3\sigma(I)]$	2.0 (2)%
$wR[F^2]$ $[I \geq 3\sigma(I)]$	3.8 (4)%
GoF $[I \geq 3\sigma(I)]$	1.2 (2)%

About 40 years ago, Abrahams *et al.* (1970) investigated a set of 17 independently measured structures of D(+)-tartaric acid. They concluded that the average R-factor error, with respect to different measurements, was about 5.8%. In our case, because of the great advances made in the development of data-collection techniques and devices, it is much improved (Table 3). The standard deviations reach values of about 0.5% [*i.e.*, an order of magnitude better than in the Abrahams *et al.* (1970) study]. Similar results are observed for the goodness-of-fit (GoF) parameter, where the mean value is equal to 1.2 (2). In the case of two data sets out of 13, the GoF is lower than 1.0, which may suggest slight over-fitting of the model. However, on the basis of the above statistics, such values of GoF are still within the 3σ significance level.

Some important information may also be extracted from the extremal residual density values computed for the whole unit cell. The average minimum value equals -0.17 (3) e Å^{-3}, whereas the average maximum value amounts to $+0.15$ (3) e Å^{-3}. This suggests that the overall mean residual density is very flat. The flatness criterion seems to be best fulfilled by the oxa12 model ($\varrho_{res}^{min/max} = -0.11/+0.12$ e Å^{-3}). The largest extreme values are observed for oxa6 ($\varrho_{res}^{min/max} = -0.22/+0.20$ e Å^{-3}); however, they are still commonly acceptable.

Additional insight into the residual density distribution is achieved by analysing the two-dimensional plots in the oxalic acid plane (Fig. 3). The map of the mean residual density values is very flat and almost featureless. On the other hand, the corresponding standard deviation map shows some non-uniform features, located mostly in the vicinity of the molecular fragments. It has to be noted that while summing up all maps some systematic errors start to appear in the regions not well modelled by the atomic multipole expansion (*i.e.*, far from atomic positions) (see supplementary material).

Finally, the so-called residual density analysis introduced by Meindl & Henn (2008) was performed and the results are summarized in the supplementary material. Most of the plots generated for the analysed models indicate Gaussian noise for the corresponding residual density maps. Only in a few cases are some asymmetric features observed, suggesting some contribution of systematic errors. These are, though, still acceptable by means of the common criteria applied in the experimental charge-density studies.

The discussion would be incomplete without a brief mention of the static deformation density maps. Such maps constitute a useful tool to analyse the quality of the multipole model fit, and in some cases they also provide additional

chemical information. The two-dimensional maps of the mean static deformation density and its standard deviation are presented in Fig. 4. There is a good reproducibility of the qualitative features in the density accumulation (*i.e.*, lone pairs and bonding density) and reduction regions. However, in the standard deviation map it is clearly seen that the largest errors are present in the vicinity of atomic positions. At the nuclei they reach values up to about 1.7 e Å^{-3}. This results from the well known problem regarding the multipole model inflexibility and hence its inability to describe the nucleus region properly (Abramov *et al.*, 2000; Spackman & Byrom, 1996;

Madsen *et al.*, 2004; Volkov *et al.*, 2001). On the other hand, much smaller maxima are observed in the chemically important areas such as the oxygen-atom lone electron pairs, suggesting that in such cases quantitative information can be obtained and binding conclusions drawn. Importantly, the bonding regions (*e.g.* hydrogen-bond interaction lines) exhibit error values smaller than 0.05 e Å^{-3}.

It is also interesting and valuable to point out the Shapiro–Wilk test performed at each point of the grid. Such a 0–1 grid (1 means that the test is *not* fulfilled, 0 otherwise) is then multiplied by the original mean grid (Figs. 3*a* and 4*a*, respectively). Such plots are presented in Fig. 5. Clearly, the normality test is well fulfilled for the residual density in the plane of the molecule. This is not the case for the deformation density. Plot contours indicate that the C1—O2 bond density and lone pair regions are those most affected by systematic errors. This agrees well with the already mentioned discussion related to Fig. 4*b*. However, the nuclei positions are not much affected. Such results may be related to the different source of errors affecting the bonding and nuclei density regions.

To test the significance of the residual density peaks on the residual map we used the Student's *t*-test, with an assumption that the normal distribution is fulfilled for the residual density. The one-sample Student's *t*-test, with a null hypothesis that the residual density should be equal to zero, was performed on every point of the residual density map (Fig. 5*c*). As expected, most of the peaks were removed and the map is clearer. However, even at the significance level of $\alpha = 0.001$, some residual density peaks are still present. Therefore, we can conclude that, at these peak positions, the residual density is
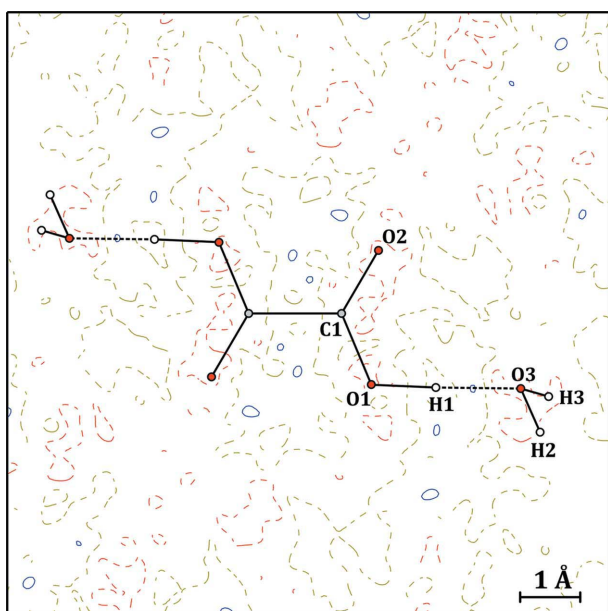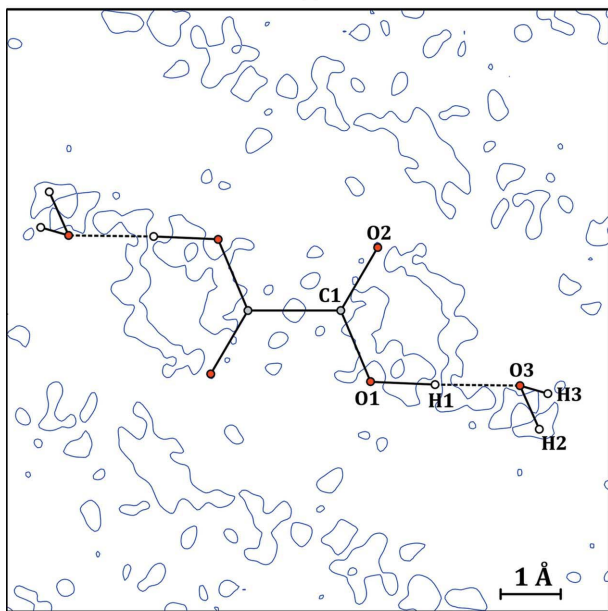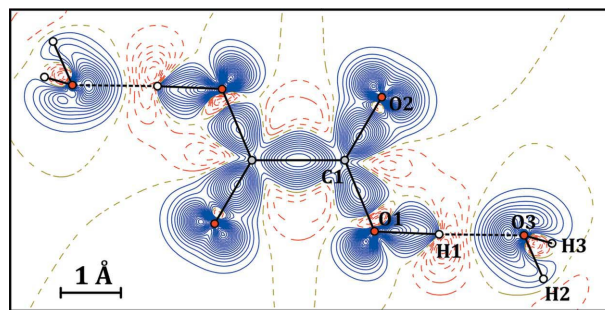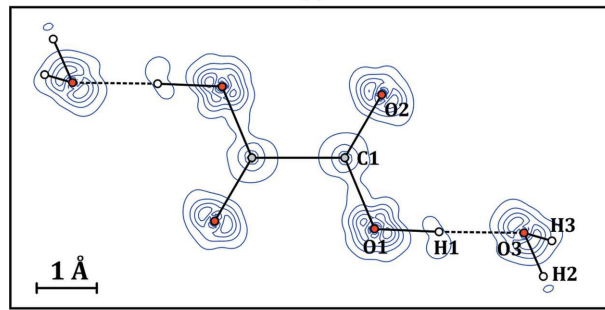

Figure 3
Two-dimensional maps of (*a*) mean residual density distribution ($\varrho_{\text{res}}^{\text{min/max}} = -0.10/+0.06$ e Å^{-3}) and (*b*) computed residual density standard deviation ($s_{\varrho_{\text{res}}}^{\text{max}} = 0.07$ e Å^{-3}) [C1—O2—C1$_{(1-x,\ 1-y,\ 1-z)}$ plane; contours at 0.05 e Å^{-3}, blue lines – positive values, red – negative].

Figure 4
Two-dimensional maps of (*a*) mean static deformation density distribution ($\varrho_{\text{def}}^{\text{min/max}} = -0.39/+1.10$ e Å^{-3}) and (*b*) its computed standard deviation ($s_{\varrho_{\text{def}}}^{\text{max}} = 1.73$ e Å^{-3}) [C1—O2—C1$_{(1-x,\ 1-y,\ 1-z)}$ plane; contours at 0.05 e Å^{-3}, blue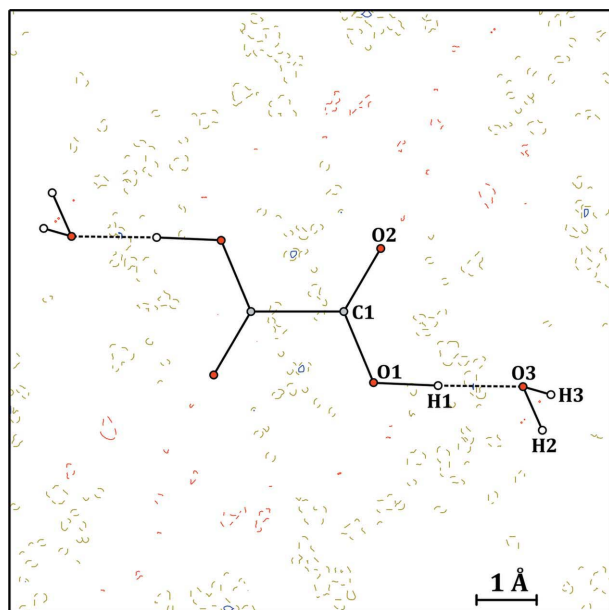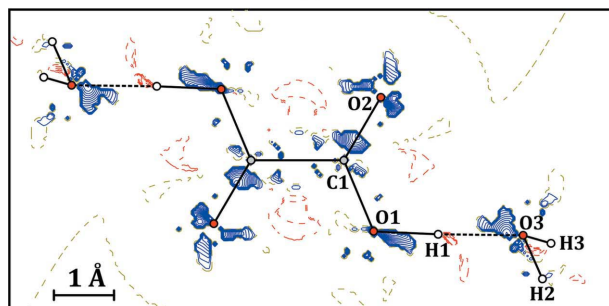 lines – positive values, red – negative]. The largest maxima in the standard deviation map are located at the nuclei positions.
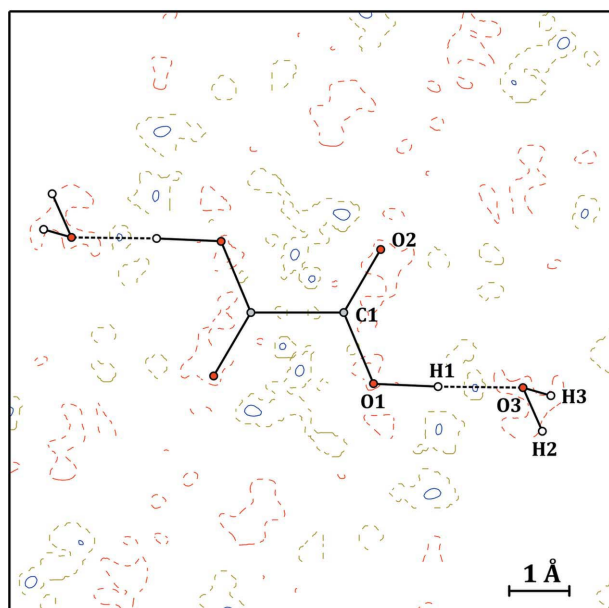
Figure 5
Two-dimensional maps of the Shapiro–Wilk test performed for (*a*) residual density and (*b*) deformation density maps [C1—O2—C1$_{(1-x, 1-y, 1-z)}$ plane; contours at 0.05 e Å^{-3}, blue lines – positive values, red – negative]. (*c*) Two-dimensional map of the mean residual density after application of the Student's *t*-test.

Table 4
Selected sample mean bond lengths, angles and hydrogen-bond parameters, with the corresponding sample standard deviations.

Bond	d_{A-B} (Å)	Angle	θ_{A-B-C} (°)
C1—C1^i	1.5427 (8)	O2—C1—C1^i	120.41 (4)
C1—O1	1.2872 (8)	O1—C1—C1^i	112.68 (4)
C1—O2	1.2223 (8)	O2—C1—O1	126.91 (4)
		C1—O1—H1	115.6 (9)
		H2—O3—H3	104 (3)
Distance	$d_{D\cdots A}$ (Å)	Hydrogen-bond angle	$\theta_{D-H\cdots A}$ (°)
H1$\cdots$O3	1.414 (2)	O1—H1$\cdots$O3	176 (2)
O1$\cdots$O3	2.486 (2)	O3—H2$\cdots$O2^{ii}	166 (3)
H2$\cdots$O2^{ii}	1.871 (9)	O3—H3$\cdots$O2^{iii}	153 (7)
O3$\cdots$O2^{ii}	2.824 (2)		
H3$\cdots$O2^{iii}	1.93 (3)		
O3$\cdots$O2^{iii}	2.827 (2)		

Symmetry transformations: (i) $-x + 1, -y + 1, -z + 1$; (ii) $x, -y + 1.5, z - 0.5$; (iii) $-x, y - 0.5, -z + 0.5$.

significantly different than zero, which suggests that there are some systematic errors in the derived charge-density models.

4.4. Geometry and thermal motion

Structural parameters which can be readily analysed are bond lengths and bond angles. The corresponding mean values and the standard deviations characterizing the investigated sample are presented in Table 4. Interestingly, the calculated standard deviations are usually very similar to the ones computed by the least-square procedure after the last refinement cycle. This shows a rather narrow distribution of these parameters and, in consequence, a very good reproducibility of geometrical results. For covalently bound atoms, variations of bond lengths are of the order of 10^{-4}, whereas for angles they are of the order of 10^{-2}. C—C and C—O bond lengths agree with the neutron diffraction results within 3σ (taking into account neutron data inaccuracies as well). The calculated mean values seem to be slightly smaller, *i.e.* of some parts of 10^{-3}, than the sxd1 model values. Nevertheless, the agreement is quite good. The obtained results imply that the least-squares estimates of bond lengths and bond angles may be considered as good estimators of the structure determination precision. It may also be concluded that for the standard low-resolution data of good quality such estimates would be an order of magnitude larger. It should be recalled here that the angles computed for bonds with hydrogen atoms are determined with lower precision than those for heavy-atom bonds. Here, the H2—O2—H3 and all O—H$\cdots$O angles show variations of several degrees (*i.e.*, the standard deviation ranges from 2 to 7°). This is rather discouraging in terms of analysis of the very fine details of hydrogen bonding. However, the multipole model still provides a significant improvement when compared to the independent-atom model (IAM) results (Hoser *et al.*, 2009). In addition, such bond-angle values and their sample standard deviations compare well with the neutron diffraction results within the 3σ limit. The same is true for a comparison of our geometrical results with the ones obtained by Martin & Pinkerton (1998) (the 3σ limit is generally fulfilled). It might

Table 5
Root-mean-square displacement factors along the principal axes for non-hydrogen atoms.

Sample means and standard deviations are given in bold. Estimated ellipsoid orientation variations are given in degrees.

| | Atomic root-mean-square displacements (Å) | | | | | |
| | O1 | | | O2 | | |
Model	u_1	u_2	u_3	u_1	u_2	u_3
oxa1	0.146	0.101	0.085	0.144	0.089	0.096
oxa2	0.152	0.106	0.090	0.151	0.093	0.101
oxa3	0.147	0.101	0.087	0.146	0.088	0.098
oxa4	0.143	0.099	0.084	0.142	0.087	0.095
oxa5	0.150	0.104	0.087	0.148	0.090	0.100
oxa6	0.147	0.101	0.086	0.146	0.089	0.097
oxa7	0.149	0.103	0.087	0.148	0.090	0.099
oxa8	0.148	0.100	0.088	0.146	0.088	0.098
oxa9	0.146	0.101	0.086	0.145	0.089	0.097
oxa10	0.149	0.105	0.088	0.147	0.091	0.101
oxa11	0.146	0.101	0.085	0.145	0.088	0.097
oxa12	0.146	0.101	0.087	0.145	0.089	0.099
oxa13	0.149	0.102	0.087	0.148	0.089	0.098
Mean values	**0.148 (2)**	**0.102 (2)**	**0.087 (2)**	**0.146 (2)**	**0.089 (2)**	**0.098 (2)**
sxd1	0.144	0.087	0.068	0.147	0.067	0.084
Direction variability	1.3 (10)			1.9 (17)		

| | Atomic root-mean-square displacements (Å) | | | | | |
| | C1 | | | O3 | | |
Model	u_1	u_2	u_3	u_1	u_2	u_3
oxa1	0.110	0.091	0.086	0.136	0.105	0.090
oxa2	0.118	0.099	0.091	0.142	0.109	0.094
oxa3	0.112	0.092	0.088	0.137	0.105	0.091
oxa4	0.110	0.090	0.085	0.134	0.103	0.089
oxa5	0.114	0.094	0.087	0.14	0.108	0.092
oxa6	0.111	0.092	0.087	0.137	0.105	0.091
oxa7	0.113	0.094	0.088	0.140	0.107	0.092
oxa8	0.111	0.092	0.086	0.138	0.106	0.091
oxa9	0.110	0.092	0.086	0.136	0.105	0.091
oxa10	0.112	0.096	0.091	0.138	0.108	0.094
oxa11	0.110	0.091	0.086	0.137	0.104	0.090
oxa12	0.111	0.092	0.087	0.136	0.106	0.092
oxa13	0.112	0.093	0.087	0.139	0.106	0.092
Mean values	**0.112 (2)**	**0.093 (2)**	**0.087 (2)**	**0.138 (2)**	**0.106 (2)**	**0.091 (1)**
sxd1	0.115	0.067	0.074	0.137	0.080	0.088
Direction variability	4.3 (36)			1.0 (4)		

be also estimated that the bond-length precision error in the case of the old data is roughly an order of magnitude larger than that resulting from our experiments (*i.e.*, the positional parameter discrepancies are up to 0.002 Å, which may indicate that the bond standard deviation is even larger).

In order to compare thermal motion parameters we computed the root-mean-square displacement factors along the principal axes for all non-hydrogen atoms. These are obtained from U_{ij}'s by well known transformations (Waser, 1955; Busing & Levy, 1958; Sands, 1982). Respective values for

all models are summarized in Table 5. Such an approach enables comparison between the probability ellipsoid 'sizes'. Almost all principal components exhibit root-mean-square displacement sample deviations of around 0.002 Å (corresponding to 4×10^{-6} Å of mean displacement) in each direction. The relative error when estimating the physical atomic displacements is then about 2% or less which is due to the thermal motion. This is a very encouraging result regarding the analysis of temperature dependence of various structural parameters (*e.g.* during the phase transition). For instance, in the case of the Laue-method-based photo-crystallographic techniques, the observed structural changes are much larger (of the order of 10^{-1} Å) (Benedict *et al.*, 2011; Makal *et al.*, 2011, 2012) and, thus, in view of the above results, they cannot be associated solely with the machine precision (*i.e.*, they seem to be a real physical effect). Similarly, well described ADPs would constitute a firm basis for reliable solid-state entropy determination from the X-ray diffraction data (Madsen & Larsen, 2007). In addition, the Shapiro–Wilk test is well fulfilled by all principal components of oxygen atoms, whereas for the carbon atom the normality is questionable. This is also reflected in the overall ellipsoid orientation variability. The estimated deviations from the average value do not exceed 2° (Table 5), except for the C1 atom where the deviation reaches about 4°. It is worth noting that all these values are not considered significantly different from zero, which suggests that there is almost no orientation change for vibrating heavy atoms. This result can be compared to the change in the ellipsoid orientation due to the temperature increase, which reached about 8° in the case of the previously studied Zn complex (Schmøkel *et al.*, 2010). Thus, it may be concluded that the ellipsoid orientation change is rather a subtle effect and should be studied with particular care.

The difference between the X-ray and neutron data also seems important here. Clearly, the corresponding ellipsoid sizes are different (Table 5). This observation justifies the scaling of hydrogen-atom ADPs, mentioned while describing the multipole refinements. Comparison of these results with the ones previously published is hampered due to the unavailability of a full set of thermal parameters in the literature.

4.5. Multipolar parameters

In a standard procedure of the data treatment unit-cell parameters are kept fixed while various other parameters are refined. These are usually a scale factor (k_S), atomic coordinates (x, y, z), thermal parameters (U_{ij}'s), and, finally, multipolar terms describing the charge redistribution (P_v and κ) and atomic deformations (P_{lm}'s). Obviously, the accuracy of these parameters is mostly affected by errors coming from two main sources: (*a*) cell parameters and (*b*) structure-factor values. It is, however, convenient to treat the refined parameters as *independent* and *random* variables in the statistical analyses. In this section, particular attention is paid to the comparative analysis of multipolar terms. The corresponding numerical values are summarized in Table 6.

Table 6
Calculated sample mean and sample standard deviation values for all refineable charge-density parameters across the considered data sets.

A *possible* local symmetry is given for comparison.

Atom	Local symmetry	Multipole populations						
		$\bar{\kappa}$	$\bar{P}_v$	$\bar{P}_{10}$	$\bar{P}_{11+}$	$\bar{P}_{11-}$	$\bar{P}_{20}$	$\bar{P}_{21+}$
O1	m	0.993 (7)	6.186 (51)	0.002 (7)	−0.054 (5)	−0.045 (7)	−0.008 (12)	−0.004 (19)
O2	m	0.993 (8)	6.188 (42)	−0.009 (10)	−0.071 (10)	−0.007 (8)	−0.065 (17)	−0.004 (12)
C1	m	0.998 (21)	4.054 (99)	0.004 (21)	0.116 (22)	0.020 (18)	−0.312 (16)	−0.006 (13)
O3	$mm2$	0.989 (8)	6.289 (61)	−0.069 (6)	0.002 (9)	−0.001 (8)	−0.001 (9)	0.005 (15)
H1	∞		0.629 (28)	0.086 (6)	−0.001 (15)	0.043 (14)	0.110 (15)	−0.001 (11)
H2	∞		0.856 (32)	0.158 (11)	−0.006 (27)	−0.006 (35)	0.094 (17)	−0.005 (15)
H3	∞		0.799 (25)	0.162 (11)	−0.043 (65)	−0.010 (40)	0.072 (25)	−0.011 (34)
		$\bar{P}_{21-}$	$\bar{P}_{22+}$	$\bar{P}_{22-}$	$\bar{P}_{30}$	$\bar{P}_{31+}$	$\bar{P}_{31-}$	$\bar{P}_{32+}$
O1	m	0.002 (8)	−0.039 (13)	0.011 (15)	−0.010 (8)	0.016 (6)	−0.018 (5)	0.0004 (52)
O2	m	0.005 (9)	−0.045 (10)	−0.005 (10)	−0.008 (6)	0.005 (7)	−0.005 (4)	−0.003 (3)
C1	m	−0.009 (13)	0.106 (19)	−0.048 (18)	0.009 (10)	0.007 (9)	−0.022 (10)	0.006 (14)
O3	$mm2$	−0.004 (18)	0.052 (7)	0.016 (24)	−0.040 (5)	0.011 (4)	0.008 (6)	−0.036 (12)
H1	∞	0.018 (12)	0.002 (13)	−0.0005 (99)				
H2	∞	0.008 (21)	−0.002 (7)	−0.003 (9)				
H3	∞	−0.013 (25)	0.037 (23)	0.005 (14)				
		$\bar{P}_{32-}$	$\bar{P}_{33+}$	$\bar{P}_{33-}$	$\bar{P}_{40}$	$\bar{P}_{41+}$	$\bar{P}_{41-}$	$\bar{P}_{42+}$
O1	m	0.002 (4)	0.053 (5)	0.010 (5)	0.005 (8)	−0.002 (9)	−0.005 (7)	−0.001 (6)
O2	m	0.002 (5)	0.031 (3)	0.003 (4)	−0.005 (11)	0.001 (9)	0.0003 (68)	−0.001 (10)
C1	m	−0.002 (15)	0.474 (17)	−0.060 (15)	0.081 (17)	0.007 (17)	−0.007 (10)	0.007 (22)
O3	$mm2$	−0.012 (9)	0.016 (8)	−0.009 (9)	−0.011 (6)	−0.003 (9)	0.007 (10)	−0.005 (12)
		$\bar{P}_{42-}$	$\bar{P}_{43+}$	$\bar{P}_{43-}$	$\bar{P}_{44+}$	$\bar{P}_{44-}$		
O1	m	0.010 (8)	0.005 (7)	−0.004 (9)	0.011 (13)	0.014 (8)		
O2	m	0.002 (5)	−0.002 (8)	−0.001 (10)	0.008 (8)	−0.001 (9)		
C1	m	−0.001 (18)	−0.015 (24)	−0.008 (16)	−0.032 (20)	0.022 (21)		
O3	$mm2$	0.002 (10)	−0.003 (8)	−0.005 (9)	0.009 (5)	0.009 (9)		

Standard deviation magnitudes of the P_v values seem to be quite large. However, their *relative* values are not that significant (*circa* 2% on average) when compared to the P_v population magnitudes. On the other hand, the same standard deviations are much more substantial when related to the net charges (*i.e.*, $Z − P_v$) as they amount to about 42% of their values. As a consequence, it is questionable to draw any further conclusions regarding, for example, charge-transfer concepts on the basis of net charges only. On the other hand, from our experience, net charge values may constitute valuable parameters when selecting a proper form factor so as to describe a heavy atom (Domagała *et al.*, 2009; Kamiński *et al.*, 2011).

The hexadecapole populations proved to be an interesting subject of study. Almost all of them are statistically insignificant and equal to zero, except for the P_{40} of the C1 atom (the parameter value is greater than 4σ). This might suggest that in most of the refinements these parameters could be left unrefined. However, having the proper data-to-parameter ratio extending the number of refined parameters with the non-statistically supported ones (or simply taking the whole expansion up to hexadecapoles) should not be a severe problem. The R_{free} approach, recently introduced into charge-density analysis by Zarychta *et al.* (2011) and Paul *et al.* (2011), would be a very useful tool to be utilized here. However, the idea of rejecting some part of the reflections during the refinement of charge density is questionable. This is due to the significant influence of various reflections on the fine details of the electron-density distribution. A more appropriate approach would be to determine a 'rigid' set of reflections which cannot be taken as a test set in the R_{free} refinement. Nevertheless, the original R_{free} procedure seems to be working quite well in the determination of how loose the charge-density model is in respect to the data set given.

It is, though, worth noting that the above result is not fully representative and, thus, cannot serve as a basis for some general refinement strategy for other structures. Oxalic acid and water molecules present in the analysed crystal structure are both quite 'flat' entities. Therefore, the hexadecapole deformation functions should not be significantly populated. However, in other structures these fine details can be obviously pronounced, especially in the case of heavier atoms located in a more complex environment (*e.g.* transition metals).

Analysis of local-symmetry-forbidden multipole populations constitutes another important aspect. *Possible* local symmetries, assigned *initially* by the *LSDB* program, are given in the second column of Table 6. Consequently, the following constraints would be imposed on the multipolar parameters (when considered up to octupoles): (i) $P_{11+} = P_{11-} = P_{21+} = P_{21-} = P_{22-} = P_{31+} = P_{31-} = P_{32-} = P_{33+} = P_{33-} = 0$ (*mm2* symmetry), (ii) $P_{10} = P_{21+} = P_{31-} = P_{32-} = 0$ (*m* symmetry) and (iii) $P_{11+} = P_{11-} = P_{21+} = P_{22+} = P_{22-} = 0$ (∞ symmetry for hydrogen atoms). As indicated by Table 6, most of these values are statistically insignificant and can be considered as equal to zero (most values do not even exceed the 1σ cutoff).

Once again, comparison of our results with those presented by Coppens *et al.* (1984) and Martin & Pinkerton (1998) sheds some light on electron-density distribution. In the latter paper, Martin & Pinkerton (1998) found a striking difference between the P_v parameters for the C1 atom ($P_v = 4.31$ for the point-detector study and $P_v = 3.87$ for the CCD). Our results indicate that within the range of 3σ (*i.e.*, 4.054 ± 0.279) both values are acceptable. Indeed, as stated by Martin & Pinkerton (1998), such a discrepancy does not affect the total electron-density distribution significantly.

4.6. Point topological descriptors

Analysis of topological properties of the electron-density distribution is one of the most common undertaken in charge-
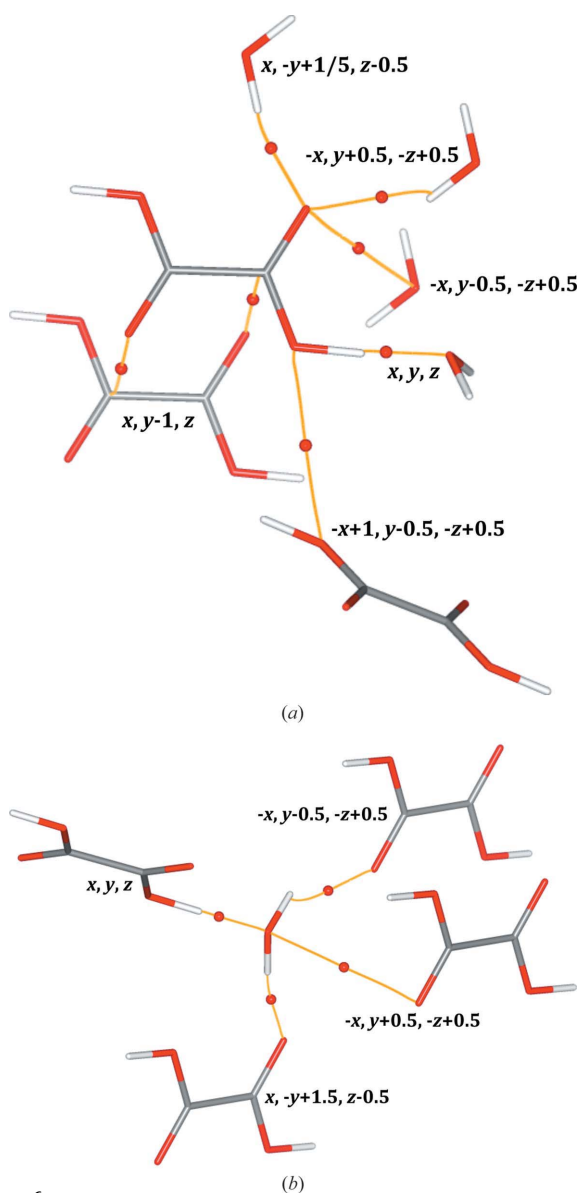

(*a*)

(*b*)

Figure 6
Bond paths (thick golden lines) and bond critical points (red spheres) present in all models of oxalic acid dihydrate: oxalic acid (*a*) and water (*b*) molecule environments. For oxalic acid moieties symmetry transformations are provided only for half of the molecule.

density studies. In most cases, various descriptors are computed at bond critical points and provide a firm basis for analysis of chemical bonding properties. It is then extremely important to analyse both the variability of such properties and the stability of a molecular graph. Here, the term 'molecular graph' is understood as a set of *all* non-redundant bond paths present in the crystalline framework (this includes both strong bonds and weak intermolecular contacts).

During the analysis of bond paths for all 13 models we found that the covalent electron-density topology is highly reproducible and all the bond paths assigned to strong covalent interactions are always present. However, this is not the case for some of the weak intermolecular contacts. Fig. 6 shows the molecular graph of the oxalic acid and water molecules in the crystal. Golden lines indicate the bond paths which are present in all cases and thus constituted the subject for further statistical analysis.

All bond critical point properties are summarized in Table 7. The most important ones are the electron density and its Laplacian. In addition, the kinetic and potential energy densities, and ellipticity values are also listed. The electron-density standard deviations constitute an important quantity in the analysis of interactions. Here, for strong covalent bonds they are in the order of 10^{-2}, whereas for weaker contacts they are even an order of magnitude smaller. Judging from Table 7 one can see that the relative error for weak contacts is also smaller. Importantly, all electron-density values show a small spread and are well determined. A similar observation is made for the Laplacian, although some more exceptions are present in this case. For example, the very strong O3$\cdots$H1 hydrogen bond exhibits the Laplacian value of $-1\,(1)\,\mathrm{e}\,\text{Å}^{-5}$. This indicates the statistical insignificance of this quantity (*i.e.*, it is not straightforward if it is positive or negative). However, a value close to zero might indicate that the electrostatic character of this contact is counterbalanced by the covalent contribution (the hydrogen bond is very short, Table 4). Similar behaviour is observed for the two remaining hydrogen bonds as well.

The sample mean and sample standard deviation of electron density and its Laplacian can also be computed *along* the bond path. Fig. 7 shows an exemplary case – the C1—O1 bond. As expected, the smallest error is observed at the BCP position and increases as we move away from the BCP. The same behaviour is observed for the relative error values, although it is much less pronounced (see supplementary material).

Other properties derived from the electron density and its Laplacian values are the kinetic and potential energy densities, which can be computed using the Abramov equation (Abramov, 1997). The Abramov assumption is relevant for the internuclear region analysis (closed-shell and shared-shell interatomic interactions) and is particularly important for the chemical bond studies. The potential energy density value at the hydrogen-bond BCP obtained *via* this approach may be utilized further for the hydrogen-bond energy estimation by means of the Espinosa–Lecomte approach (Espinosa *et al.*, 1998, 1999). Therefore, we checked the variability of the kinetic and potential energy density values for the covalent

Table 7
Mean values and standard deviations of selected properties evaluated at bond critical points of covalent bonds and intermolecular contacts (ϱ, electron density; $\nabla^2\varrho$, Laplacian of electron density; G, kinetic energy density; V, potential energy density; ε, ellipticity).

Property at BCP	Covalent bonds					
	C1—O1	C1—O2	C1—C1^i	O1—H1	O3—H2	O3—H3
ϱ (e Å^{-3})	2.50 (6)	2.91 (6)	1.75 (3)	1.80 (5)	2.67 (5)	2.60 (4)
$\nabla^2\varrho$ (e Å^{-5})	−29 (4)	−25 (7)	−15 (1)	−38 (3)	−58 (5)	−56 (6)
G (kJ mol^{-1} a$_0^{-3}$)	919 (62)	1408 (80)	523 (12)	145 (60)	550 (79)	524 (97)
V (kJ mol^{-1} a$_0^{-3}$)	−2616 (90)	−3494 (67)	−1452 (34)	−1323 (75)	−2685 (96)	−2573 (72)
ε	0.10 (4)	0.10 (4)	0.16 (4)	0.03 (3)	0.02 (1)	0.09 (4)

Property at BCP	Intermolecular contacts					
	O3···H1	O1···O1^{ii}	C1···O2^{iii}	O2···O3^{iv}	O2···H3^v	O2···H2^{vi}
ϱ (e Å^{-3})	0.61 (3)	0.026 (1)	0.054 (4)	0.025 (1)	0.17 (2)	0.21 (2)
$\nabla^2\varrho$ (e Å^{-5})	−1 (1)	0.47 (2)	0.61 (3)	0.43 (1)	1.5 (6)	1.2 (3)
G (kJ mol^{-1} a$_0^{-3}$)	117 (14)	9.2 (3)	13.5 (8)	8.5 (3)	44 (9)	45 (5)
V (kJ mol^{-1} a$_0^{-3}$)	−264 (17)	−5.7 (3)	−10.4 (8)	−5.3 (2)	−46 (4)	−57 (6)
ε	0.08 (6)	0.4 (5)	0.3 (2)	0.8 (4)	0.3 (4)	0.07 (6)

Symmetry transformations: (i) $-x + 1, -y + 1, -z + 1$; (ii) $-x + 1, y - 0.5, -z + 0.5$; (iii) $x, y - 1, z$; (iv) $-x, y - 0.5, -z + 0.5$; (v) $-x, y + 0.5, -z + 0.5$; (vi) $x, -y + 1.5, z + 0.5$.

bonds and selected interatomic contacts across the 13 charge-density models. It appeared that all the energy density values are statistically significant with the average relative error reaching about 10% (with the highest value for the kinetic energy density characterizing the O1—H1 bond BCP equal to 41%). The relative sample error is generally noticeably lower for the potential energy density than that for the kinetic energy density for both covalent bonds and intermolecular BCPs. Keeping in mind the above considerations, it can be concluded that the energy density results are consistent, among which especially the potential energy density values are essential for hydrogen-bond energy estimation. Furthermore, based on the Espinosa–Lecomte formula and our results (Table 7), the estimated hydrogen-bond energy uncertainty (denoted by sample standard deviation) does not exceed 9 kJ mol^{-1} in the case of the selected intermolecular contacts, whereas the relative error does not constitute more than 11% of the mean energy value, which reflects the precision of the derived energies. This makes the analysis of interaction energies, by the use of the Espinosa–Lecomte approach, feasible.

The last parameter, the ellipticity, seems to be quite sensible. For stronger interactions its values are statistically significant at the 1σ level, though at 3σ it might be questionable in some cases (e.g., the C1—O1 bond). For weak intermolecular contacts, ellipticity is, as might be expected, rather poorly defined.

4.7. Integrated properties

Integrated atomic charges and atomic volumes are all very important quantities which can be derived directly from the electron-density distribution. Integrated atomic charges are very frequently used in the analysis of various effects associated with charge redistribution under weak interaction

formation, molecule reactivity etc. The numerical values are shown in Table 8. The estimated standard deviations reach values of 0.1 e. The average ratio of standard deviation over the corresponding mean charge amounts to about 5%. In our opinion, this is in the range of possibly observed charge differences between different atoms in the same molecule. Thus, one needs to be very careful to draw meaningful conclusions on the basis of atomic charges obtained experimentally. For instance, in the case of oxalic acid data, the O1 and O2 atoms have practically the same charge (the difference is statistically insignificant), yet they are obviously chemically different. The same thing is true for atomic volumes. In this case the variations are in the range from 1 to 10% even.
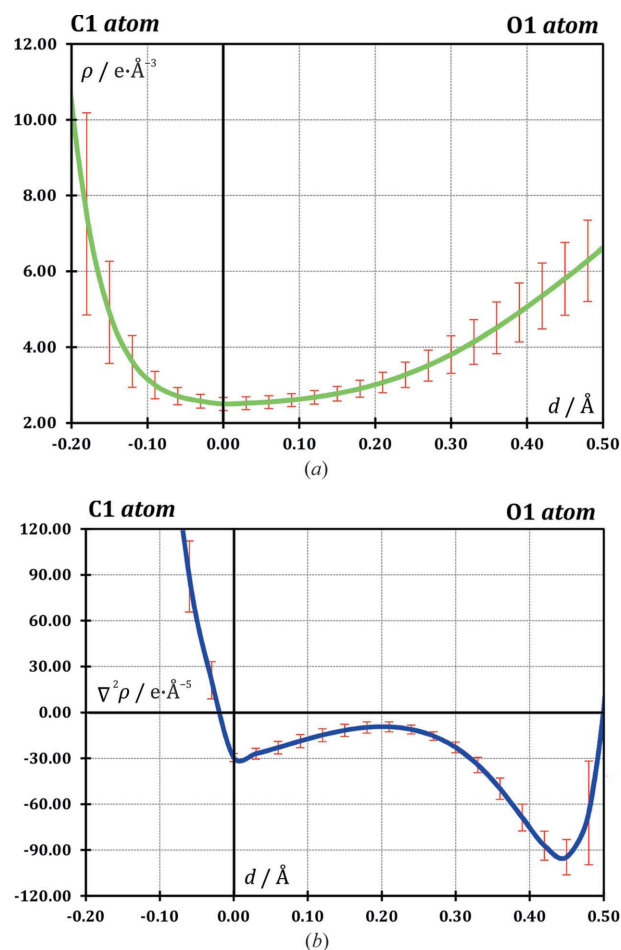

Figure 7
Sample means and sample standard deviations evaluated at each point of the C1—O1 bond path for (a) electron density (green solid line) and (b) its Laplacian (blue solid line). Error bars (red vertical lines) are shown at the 3σ and 1σ levels for (a) and (b) panels, respectively.

Radosław Kamiński et al. · Statistical analysis of structural parameters

research papers

Table 8
Sample mean values and sample standard deviations of integrated charges and integrated atomic basin volumes.

Atom	$\overline{Q}$ (e)	$\lvert s_Q/\overline{Q}\rvert$ (%)	$\overline{V}$ (Å^3)	$\lvert s_V/\overline{V}\rvert$ (%)
O1	−1.12 (7)	6.29	15.90 (17)	1.07
O2	−1.22 (8)	6.89	16.60 (30)	1.80
C1	+1.60 (10)	6.31	5.09 (23)	4.49
H1	+0.72 (2)	3.00	0.93 (9)	9.93
O3	−1.14 (5)	4.60	18.31 (18)	1.00
H2	+0.57 (3)	5.17	2.04 (17)	8.57
H3	+0.60 (2)	3.92	2.12 (16)	7.61

Table 9
Specific atoms (Q_i, i = O3, H2 and H3) and total water molecule (Q_w) charges.

Model	Q_{O3} (e)	Q_{H2} (e)	Q_{H3} (e)	Q_w (e)
oxa1	−1.17	+0.56	+0.58	−0.04
oxa2	−1.07	+0.58	+0.63	+0.14
oxa3	−1.06	+0.55	+0.63	+0.12
oxa4	−1.22	+0.62	+0.64	+0.03
oxa5	−1.10	+0.55	+0.58	+0.03
oxa6	−1.17	+0.57	+0.60	+0.01
oxa7	−1.07	+0.53	+0.58	+0.04
oxa8	−1.17	+0.57	+0.61	+0.01
oxa9	−1.19	+0.62	+0.60	+0.02
oxa10	−1.18	+0.59	+0.58	−0.01
oxa11	−1.13	+0.53	+0.57	−0.04
oxa12	−1.09	+0.56	+0.60	+0.06
oxa13	−1.14	+0.55	+0.59	+0.00
Average charges	−1.14 (5)†	+0.57 (3)†	+0.60 (2)†	**+0.03 (5)**

† Values taken from Table 8.

At this point it is worth having a more careful look at the total *molecular* charges. To simplify the discussion the attention is focused solely on the water molecule. The results are summarized in Table 9. The total charge of the water molecule, for the ith data set, is obviously computed as

$$Q_{w,i} = Q_{O3,i} + Q_{H2,i} + Q_{H3,i}$$

and the formula stands also for the sample means. The estimator of the standard deviation is derived from the whole population of the computed water charge values (using the basic equation for the standard deviation):

$$s_{Q_w} = \left[\frac{1}{n-1} \sum_{i=1}^{n} \left(\bar{Q}_w - Q_{w,i} \right)^2 \right]^{1/2}.$$

The total charge of the water molecule is statistically equal to zero, even at the 1σ level. What is important is the standard deviation regarding the whole water moiety, which is slightly larger than the corresponding standard deviations for each individual atom, *i.e.*, O3, H2 and H3. This suggests some error accumulation while summing up the atomic charges. On the other hand, the atomic charges are *not* independent variables (their sum must be obviously equal to zero over all atoms), and so the correlation factors may play a significant role (though here not really well observed). This result implies the importance of careful analysis of such charges when it is applied to investigate problems such as 'charge transfer' or

crystallographic equivalence of molecules in the asymmetric unit. It seems that an approximation of two molecules of the same compound, despite being symmetry non-equivalent, having the same charge (*e.g.* neutral) should hold rather well. Further modelling of such fine details is within the experimental errors and should not be carried out in most cases. This is in contrast with the recent literature contributions (Nelyubina *et al.*, 2010, 2011), where such an approach is arbitrarily proposed without any deeper insight into the method limitations.

4.8. Dipole and quadrupole moments

The average value of the dipole moment for the water molecule in oxalic acid dihydrate structure equals 2.1 (3) D, being close to the Hartree–Fock theoretical calculations (2.2 D) reported by Spackman (1992). The average dipole moment magnitude is also very close to the value reported by Stevens & Coppens (1980) [2.1 (2) D] on the basis of a high-resolution charge-density study. Importantly, the Shapiro–Wilk test is fulfilled showing the normal distribution of the observed dipole moment magnitudes across different measurements. The computed dipole moments show some spatial spread around their mean position [$\theta = 12$ (6)°]. This clearly indicates that the statistical variation is of the same range to the one observed from different multipolar refinements with respect to the same data set (Poulain-Paul *et al.*, 2012; Bąk *et al.*, 2011).

Second moments of charge distribution were also calculated. The average eigenvalues of the quadrupole moment in the traceless convention are as follows: $Q_{xx} = -0.7$ (1) e Å^{-2}, $Q_{yy} = -4.6$ (4) e Å^{-2} and $Q_{zz} = +5.2$ (5) e Å^{-2}. All mean values are well above the estimated standard deviations. In addition, as might be expected, the first component is very small. It is due to the molecule planarity. Similarly to dipole moments, the Shapiro–Wilk test always indicates the normal distribution of quadrupole moment values.

4.9. Electrostatic energy

In general, the multipolar model provides some information concerning the strength of the electrostatic interactions in the crystal lattice. The electrostatic interaction energies can be simply derived from the electron-density distribution. However, to draw any binding conclusions on the basis of such results, it is important to check their reliability. Table 10 contains the electrostatic energy values of selected interactions, averaged over the 13 analysed multipolar refinements. The estimated standard deviations are given in parentheses. Additionally, the theoretical results from the *PIXEL* calculations are listed as a reference point. The calculations were performed for the model oxa8, considered to be the one least deviating in geometry from the mean value. In the case of the theoretical results, the electrostatic energy was summed up together with the polarization component, as polarization should be present in the experimental data, and, thus, to some extent accounted for by the multipolar model (Bąk *et al.*, 2012).

Table 10
Interaction energy values computed between the central acid molecule and its surroundings.

Total interaction energies are computed with the *CRYSTAL* (E_{tot}^{CRY}) [DFT(B3LYP)/pVTZ] and *PIXEL* (E_{tot}^{PIX}) programs. *PIXEL*-computed interaction energy contributions: electrostatic (E_{elst}^{PIX}), dispersion (E_{disp}^{PIX}), polarization (E_{pol}^{PIX}) and repulsion (E_{rep}^{PIX}). The experimental electrostatic energy was computed by *VMOPRO* (E_{elst}^{exp}). For oxalic acid the symmetry operator is given for half of the molecule.

Motif	E_{tot}^{CRY} (kJ mol^{-1})	E_{tot}^{PIX} (kJ mol^{-1})	E_{elst}^{PIX} (kJ mol^{-1})	E_{disp}^{PIX} (kJ mol^{-1})	E_{pol}^{PIX} (kJ mol^{-1})	E_{rep}^{PIX} (kJ mol^{-1})	$E_{elst}^{PIX} + E_{pol}^{PIX}$ (kJ mol^{-1})	E_{elst}^{exp} (kJ mol^{-1})
A	−40.0	−27.4	−129.3	−82.9	−11.4	+196.2	−140.7	−158 (9)
B	−19.0	−22.2	−38.3	−14.9	−11.4	+42.4	−49.7	−58 (16)
C	−14.5	−17.4	−27.0	−9.6	−7.4	+26.6	−34.4	−45 (17)
D	−3.5	−5.2	−2.7	−0.4	−2.3	+0.2	−5.0	−12 (9)
E	−7.5	−8.4	−5.7	−1.0	−3.5	+1.8	−9.2	−13 (9)
F	−9.6	−6.4	−3.8	−3.2	−13.7	+14.3	−17.5	+11 (19)
G	−7.3	−13.7	−11.2	−1.3	−2.9	+1.7	−14.1	−12 (5)

Assigned symmetry operators and motif type: (A) acid$\cdots$water (x, y, z); (B) acid$\cdots$water ($x, -y + 1.5, z + 0.5$); (C) acid$\cdots$water ($-x, y + 0.5, -z + 0.5$); (D) acid$\cdots$water ($-x + 1, y + 0.5, -z + 0.5$); (E) acid$\cdots$water ($-x, y - 0.5, -z + 0.5$); (F) acid$\cdots$acid ($x, y + 1, z$); (G) acid$\cdots$acid ($-x + 1, y - 0.5, -z + 0.5$).

The estimated standard deviations are rather high, varying in the range of 5–17 kJ mol^{-1}. Such a result is comparable to the root-mean-square deviation values obtained for different databases of the transferable aspherical atom models with respect to the corresponding theoretical calculations (Jarzembska & Dominiak, 2012; Bąk *et al.*, 2011; Volkov *et al.*, 2006). In the case of oxalic acid, only two electrostatic interactions are significant, *i.e.*, with the electrostatic interaction energy values above 3σ. This suggests that the electrostatic interactions derived from the multipolar model should be treated more qualitatively than quantitatively, maybe except for the considerably stronger ones.

However, the general trend of the strength of interactions is preserved when compared to the sum of the electrostatic and polarization components from *PIXEL* calculations. The theoretical values are systematically lower in magnitude, yet within the 2σ margin from the multipolar model results. The largest difference is observed for the *F* interaction being of the $\pi\cdots\pi$ type. The correlation between the theoretical model and the experimental one is high (with the linear determination coefficient, R^2, being equal to 0.978). Nevertheless, the semi-empirical results of *PIXEL* are not fully consistent with the *CRYSTAL* results when the total energy is considered. The balance sheet of all the methods perfectly illustrates the accuracy and credibility of the obtained results.

4.10. Experimental electrostatic potential

Electrostatic potential (ESP, φ) is very frequently used in the analysis of crystal packing and interactions in molecular crystals (Spackman *et al.*, 2008). For this purpose it is usually mapped onto molecular surfaces such as van der Waals or Hirshfeld surfaces. The latter ones are especially convenient as they do not interpenetrate each other, but rather adjoin one another (Spackman & Byrom, 1997*b*; Spackman & Jayatilaka, 2009). Fig. 8(*a*) shows the mean electrostatic potential mapped on the Hirshfeld surface of the oxalic acid moiety. The electrostatic potential distribution follows the expected behaviour. The hydrogen-bonded regions (*i.e.*, areas closest to the H1 and O2 atoms) show the extreme electrostatic potential values. The side of the molecule (*i.e.*, the area responsible for $\pi\cdots\pi$

interaction) seems to be slightly negatively charged. Additional information about the electrostatic potential distribution is gained by the analysis of its standard deviation. It appears that the best way to visualize the standard deviation distribution is to map the $\overline{\varphi}/s_\varphi$ quantity onto the same Hirshfeld surface (Fig. 8*b*). It provides information about the statistical significance of the electrostatic potential values at given points on the surface. It is obvious and confirmed by the data that the hydrogen-bonded regions are most significant. The vicinity of the H1 atom shows very large significance ($> 10\sigma$) of the ESP values. Similarly, the negative values are quite significant in the region associated with the O2 atom ($> 4.5\sigma$). It is worth noting that the analysis of the ESP agrees well with the BCP standard deviation summary. For hydrogen-bonded regions the standard deviations are very small and, thus, the evaluated properties are well defined. Interestingly, for the $\pi\cdots\pi$ region the electrostatic potential is statistically not significant. In this case, the standard deviations and the mean ESP values are of the same range. Thus, the analysis of weaker dispersive interactions should be undertaken with extra care. This statement agrees well with the experimental electrostatic energy values computed for the *F* dimer and with the results discussed above.

In order to characterize the ESP distribution, the surface quantities, first introduced by Murray & Politzer (Murray *et al.*, 2000; Murray & Politzer, 1998), were computed. Statistical analysis of the ESP distribution was performed on three-dimensional grids with a 0.1 Å sampling step around the whole molecule of oxalic acid with a 3 Å margin. The surface quantities were evaluated on the Hirshfeld surface and around the van der Waals (vdW) surface within a shell of ± 0.15 Å thickness. The average values together with sample standard deviations are summarized in Table 11 (for a detailed definition of all parameters see the supplementary material). A comprehensive interpretation of all the ESP parameters is given in the original papers by Murray & Politzer, whereas for the purpose of our study we shall focus on the parameter reliability and overall variability.

Most of the ESP surface quantities are generally well defined at the 3σ threshold level. The mean ESP value ($\overline{V}_S$) is equal to -0.03 (4) e Å^{-1}, which correlates well with the

Table 11
ESP quantities (Π – average deviation of the surface ESP, σ_{tot}^2 – total variance of the surface ESP, ν – degree of balance; for details see the supplementary material) computed at the Hirshfeld and vdW surfaces for the oxalic acid molecule.

The sample mean values and their sample standard deviations were computed over the 13 grids.

Surface quantity	Hirshfeld surface	van der Waals surface
$V_{S,max}$ (e Å^{-1})	0.88 (5)	0.22 (4)
$V_{S,min}$ (e Å^{-1})	−0.19 (5)	−0.19 (4)
$\overline{V}_S^+$ (e Å^{-1})	0.12 (1)	0.09 (2)
$\overline{V}_S^-$ (e Å^{-1})	−0.10 (3)	−0.10 (3)
$\overline{V}_S$ (e Å^{-1})	−0.03 (4)	−0.03 (4)
Π (e Å^{-1})	0.096 (8)	0.088 (7)
σ_+^2 (e^2 Å^{-2})	0.020 (2)	0.003 (1)
σ_-^2 (e^2 Å^{-2})	0.003 (1)	0.003 (1)
σ_{tot}^2 (e^2 Å^{-2})	0.023 (3)	0.006 (1)
ν	0.11 (2)	0.22 (3)

slightly negative charge of the oxalic acid molecule, as calculated using the QTAIM approach (Table 9). However, one needs to keep in mind that the value is close to zero and its standard deviation suggests statistical insignificance (this also agrees with the QTAIM result). The average positive ($\overline{V}_S^+$), negative ($\overline{V}_S^-$) and total ($\overline{V}_S$) electrostatic potential is comparable, no matter whether it is calculated on the Hirsh-

feld or vdW surfaces. Generally speaking, the positive potential region is naturally well defined in the vicinity of the acidic protons and described with the relatively high amplitude of the positive potential values [$V_{S,max} = +0.88$ (5) e Å^{-1}; value for the Hirshfeld surface]. In turn, the negative potential is much more uniformly spread over the remaining parts of the molecule with the pronounced maxima at the oxygen atoms [$V_{S,min} = −0.19$ (5) e Å^{-1}; value for the Hirshfeld surface], but with a larger relative error than that of the positive potential. One can observe that the maximum value of the positive potential and the variance of the positive potential (σ_+^2) is higher for the Hirshfeld surface. A larger spread of the positive and negative variances for the Hirshfeld surface leads to the significant difference in the degree of balance parameter (ν). In the case of the Hirshfeld surface ν amounts to 0.11 (2), whereas for the vdW it is equal to 0.22 (3), which approaches the theoretical limit of 0.25. The difference between the results obtained for the two surfaces might be associated with the fact that vdW surfaces interpenetrate one another in the crystal frame, and thus may not be an appropriate choice for the purpose of molecular ESP distribution investigations in the solid state.

4.11. Periodic energy computations

Finally, we decided to investigate the energetic consequences of slightly different molecular geometries resulting from the obtained 13 charge-density models. Additionally, two geometry optimizations were carried out in the *CRYSTAL* program using the periodic DFT(B3LYP) method with two all-electron basis set types, *i.e.*, pVTZ (th1) and 6-31G** (th2). In the case of all 13 geometries, the total unit-cell, molecule and cohesive energies were computed with the pVTZ basis set and analysed. However, it should be stressed that the so-called 'molecule energy' stands for the energy of a molecular unit built of one acid moiety and two water molecules (it consists of two asymmetric unit fragments related through the centre of symmetry). As a result, the so-called 'cohesive energy' was calculated with respect to the mentioned three-molecule fragment. Such a choice was dictated by technical facility and is sufficient for the purpose of this analysis.

All the results are grouped together in Table 12. The absolute energy values are given in the first three columns. As can be seen, the energy results are very much consistent. Such observation is supported by the corresponding calculated mean energies and standard deviations describing the sample. Standard deviation magnitudes are quite low, as they do not exceed 1 kJ mol^{-1} for each of the energy types (unit-cell, molecule and cohesive energy). What is more, even a 3σ margin is certainly comparable to the error bars of the computation method itself.

The other columns of Table 12 contain the relative energy values with respect to those obtained for the two optimized geometries. In all cases the unit-cell energy (E_{bulk}) is lowest in the case of the optimized models, especially for the one with the more advanced basis set used during the optimization procedure (th1). The differences in the unit-cell energies
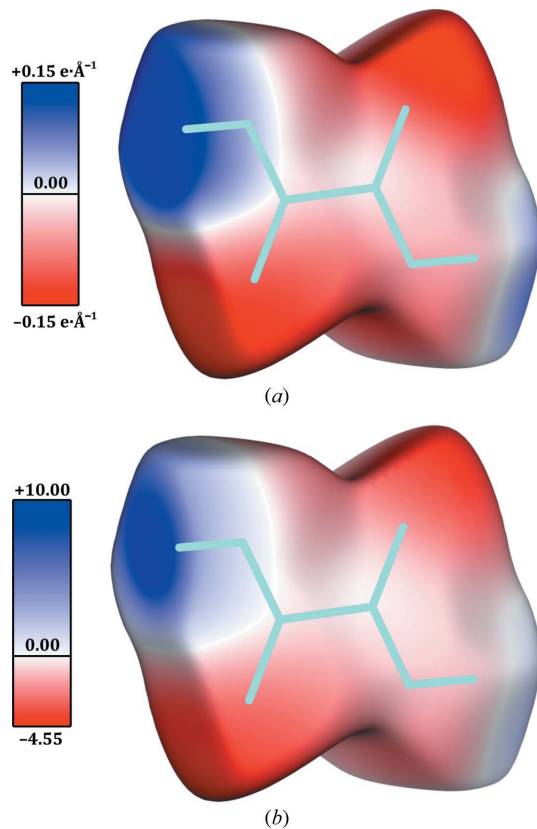
Figure 8
(a) Mean electrostatic potential ($\overline{\varphi}$) mapped onto the Hirshfeld surface of the oxalic acid molecule. (b) $\overline{\varphi}/s_\varphi$ quantity mapped onto the same Hirshfeld surface.

Table 12
Cohesive (E_{coh}), unit-cell (E_{bulk}) and molecule (E_{mol}) energies computed for all models on the basis of their geometry.

Theoretical models are included for comparison and are not used in the computation of mean values and standard deviations. $\Delta E'_i$ and $\Delta E''_i$ denote energies computed with respect to th1 and th2 models.

Model	E_{coh} (kJ mol^{-1})	E_{bulk} (kJ mol^{-1})	E_{mol} (kJ mol^{-1})	$\Delta E'_{\text{coh}}$ (kJ mol^{-1})	$\Delta E'_{\text{bulk}}$ (kJ mol^{-1})	$\Delta E'_{\text{mol}}$ (kJ mol^{-1})	$\Delta E''_{\text{coh}}$ (kJ mol^{-1})	$\Delta E''_{\text{bulk}}$ (kJ mol^{-1})	$\Delta E''_{\text{mol}}$ (kJ mol^{-1})
oxa1	−238.9	−2789600.2	−1394544.6	−5.3	15.4	12.8	1.4	13.2	5.0
oxa2	−239.8	−2789608.5	−1394547.7	−6.2	7.0	9.8	0.4	4.8	2.0
oxa3	−239.7	−2789602.2	−1394544.6	−6.1	13.4	12.8	0.6	11.1	5.0
oxa4	−242.3	−2789589.8	−1394535.8	−8.6	25.8	21.6	−2.0	23.6	13.9
oxa5	−237.4	−2789598.6	−1394545.2	−3.8	16.9	12.2	2.8	14.7	4.4
oxa6	−239.2	−2789604.2	−1394546.2	−5.6	11.4	11.2	1.0	9.2	3.4
oxa7	−240.2	−2789609.6	−1394547.7	−6.6	6.0	9.7	0.0	3.7	1.9
oxa8	−240.7	−2789610.0	−1394547.6	−7.1	5.6	9.9	−0.5	3.4	2.1
oxa9	−242.9	−2789590.1	−1394535.2	−9.3	25.4	22.2	−2.7	23.2	14.4
oxa10	−239.2	−2789610.6	−1394549.6	−5.6	5.0	7.9	1.1	2.8	0.1
oxa11	−238.1	−2789606.2	−1394548.4	−4.5	9.4	9.0	2.1	7.2	1.2
oxa12	−239.8	−2789609.5	−1394548.4	−6.2	6.1	9.1	0.4	3.8	1.3
oxa13	−240.6	−2789610.2	−1394547.8	−7.0	5.4	9.6	−0.4	3.1	1.8
th1	−233.6	−2789615.6	−1394557.4						
th2	−240.2	−2789613.3	−1394549.6						

Mean values	$\bar{E}_{\text{coh}}$ (kJ mol^{-1})	$\bar{E}_{\text{bulk}}$ (kJ mol^{-1})	$\bar{E}_{\text{mol}}$ (kJ mol^{-1})
	−239.9 (15)	−2789603.8 (73)	−1394545.3 (46)

range from 5 to almost 26 kJ mol^{-1}. On the other hand, both cohesive and molecule energies contribute to the overall stability of the crystal, and so the lowest cohesive energy may not mean the optimal molecular geometry at the same time. Most balanced and energetically closest to the results obtained for th1 and th2 is the oxa10 charge-density model, while oxa4 and oxa9 seem to be furthest away from the optimal results. However, the energetic differences constitute just a small percentage of the parent energy values.

All in all, on the basis of charge-density features it was not possible to indicate the optimal model, which is not easily justifiable on the basis of the energetic results either. Nevertheless, it should be concluded that all the energy results are very much consistent and within the error of the methods, which makes them all equally good.

5. Summary and conclusions

The results presented in this contribution can be summarized as follows:

(i) Raw intensity analysis indicates that the normality distribution is well preserved for about 60–70% of measured reflections (up to the resolution of 1.0 Å^{-1}). In the case of weak reflections the corresponding percentage is lower (around 50%).

(ii) Unit-cell parameters are determined to the order of 10^{-3} (for cell edges) and 10^{-2} (for angles). The distribution of unit-cell parameters with respect to different measurements fulfils the normality test. Values compare well with the older studies of the same compound and with the neutron diffraction data set.

(iii) Fit discrepancy factors are determined within a 0.5% range. Similarly, the standard deviation of the GoF parameter equals about 0.2. The residual density extrema are ± 0.16 (3) e Å^{-3}, which indicates a rather flat residual charge-density distribution across the whole unit cell.

(iv) Two-dimensional plots show that the mean residual density is very flat and featureless; however, its largest standard deviation values are located near molecular fragments and so it is not completely uniform. The Shapiro–Wilk test confirms the normal distribution of the experimental errors on Fourier maps.

(v) Deformation density is well reproducible between different measurements. However, the largest errors are located at and near nuclei positions. Moreover, some systematic errors (induced presumably by the multipolar model) are present, as indicated by the normality test.

(vi) Geometry is extremely well reproducible between different data sets. This is confirmed by both direct comparison and periodic computations.

(vii) The largest errors are on the valence-shell charge-transfer parameters (*i.e.*, P_v's), which are further propagated onto the net charges. The symmetry of multipolar expansions, with respect to local coordinate systems, is well preserved.

(viii) The 'always-present' bond paths are the ones related to the strong covalent bonds and well defined intermolecular contacts (such as relatively strong hydrogen bonds).

(ix) Standard deviations for electron density at bond critical points are well defined. They are especially small for well defined hydrogen-bonded contacts. The same is true for kinetic and potential energy densities, directly related to the interaction strength.

(x) Standard deviations for the integrated atomic charges are equal to about 0.1 e or similar. The charge transfer seems to be a fine detail, which needs to be taken carefully into chemical analyses.

(xi) Dipole moments are comparable with the ones presented in different, older studies. The standard deviation is small (0.2 D); however, the dipole moment direction is

Radosław Kamiński *et al.* · Statistical analysis of structural parameters **89**

less well defined [12 (6)° deviation from the mean dipole moment].

(xii) The electrostatic energies should be treated qualitatively (as suggested in the literature) as large standard deviations occur. On the other hand, they are well correlated with the *PIXEL* results.

(xiii) Similarly as for the point properties, the ESP distribution is best defined in the hydrogen-bonded regions.

(xiv) The cohesive energy values obtained for crystal geometries resulting from the studied data sets show no significant differences. Therefore, the molecular geometry is not much affected by the final model.

On the whole, experimental charge-density studies are useful in several respects. Firstly, the geometry is well determined after multipolar refinement, so one could use it as a subject for further theoretical computations. The derived properties, such as bond critical point electron densities, are extremely useful in the analysis of moderate-strength intermolecular contacts, occurring in the solid state. This conclusion may be extended to a weak non-classical bonding occurring in coordination complexes, assuring the proper description of the metal centre. Particular attention should be paid to the analysis of integrated charges and, in the case of dynamical systems (phase transitions *etc.*), unit-cell constants. Hopefully, these quantities can be much improved by the recent advances in pixel detectors. Integrated charges should be analysed with great care, especially while considering charge-transfer issues. As predicted, the very useful properties such as dipole moments, ESP distribution and electrostatic energies should be treated rather qualitatively. Of course, some quantities are affected by the multipolar model itself. Thus, it would be recommended to conduct a similar analysis using different electron-density reconstruction methods in the future, as the awareness of the model and method limitations is of prime importance. Nevertheless, we clearly show here that the source of errors is Gaussian-noise-like and it is particularly related to the experimental setup and methodology inaccuracies. It should also be noted that the variations of the derived properties will likely be larger when measured on a greater variety of instruments. We do believe that the extension of our study to different types of detectors [*e.g.* complementary metal-oxide superconductor (CMOS) or other pixel-array devices] and X-ray sources (*e.g.* synchrotron radiation) will be available in the future and the presented results can be readily updated.

References

Abrahams, S. C., Hamilton, W. C. & Mathieson, A. McL. (1970). *Acta Cryst.* A**26**, 1–18.

Abramov, Yu. A. (1997). *Acta Cryst.* A**53**, 264–272.

Abramov, Y. A., Volkov, A., Wu, G. & Coppens, P. (2000). *J. Phys. Chem. B*, **104**, 2183–2188.

Agilent Technologies (2010). *CRYSALIS PRO*. Agilent Technologies, Yarnton, Oxfordshire, UK.

Allen, F. H. (2002). *Acta Cryst.* B**58**, 380–388.

Bader, R. F. W. (1994). *Atoms in Molecules: A Quantum Theory*. New York: Oxford University Press.

Bąk, J. M., Czyżnikowska, Ż. & Dominiak, P. M. (2012). *Acta Cryst.* A**68**, 705–714.

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*a*). *J. Chem. Phys.* **88**, 2547–2553.

Becke, A. D. (1988*b*). *Phys. Rev. A*, **38**, 3098–3100.

Becker, P. J. & Coppens, P. (1974*a*). *Acta Cryst.* A**30**, 148–153.

Becker, P. J. & Coppens, P. (1974*b*). *Acta Cryst.* A**30**, 129–147.

Becker, P. J. & Coppens, P. (1975). *Acta Cryst.* A**31**, 417–425.

Benedict, J. B., Makal, A., Sokolow, J. D., Trzop, E., Scheins, S., Henning, R., Graber, T. & Coppens, P. (2011). *Chem. Commun.* **47**, 1704–1706.

Blessing, R. H. (1987). *Crystallogr. Rev.* **1**, 3–58.

Blessing, R. H. (1989). *J. Appl. Cryst.* **22**, 396–397.

Blessing, R. H. (1995*a*). *Acta Cryst.* A**51**, 33–38.

Blessing, R. H. (1995*b*). *Acta Cryst.* B**51**, 816–823.

Brandt, S. (1998). *Data Analysis: Statistical and Computational Methods for Scientists and Engineers*. London: Springer.

Bruker AXS Inc. (2013). *APEX2*. Madison, Wisconsin, USA.

Busing, W. R. & Levy, H. A. (1958). *Acta Cryst.* **11**, 450–451.

Civalleri, B., Zicovich-Wilson, C. M., Valenzano, L. & Ugliengo, P. (2008). *CrystEngComm*, **10**, 405–410.

Cochran, W. G. (1934). *Math. Proc. Camb. Philos. Soc.* **30**, 178–191.

Coppens, P. (1997). *X-ray Charge Densities and Chemical Bonding*. New York: Oxford University Press, Inc.

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**, 754–756.

Dittrich, B., Sze, E., Holstein, J. J., Hübschle, C. B. & Jayatilaka, D. (2012). *Acta Cryst.* A**68**, 435–442.

Domagała, S., Korybut-Daszkiewicz, B., Straver, L. & Woźniak, K. (2009). *Inorg. Chem.* **48**, 4010–4020.

Dovesi, R., Orlando, R., Civalleri, B., Roetti, C., Saunders, V. R. & Zicovich-Wilson, C. M. (2005). *Z. Kristallogr.* **220**, 571–573.

Dovesi, R., Saunders, V. R., Roetti, C., Orlando, R., Zicovich-Wilson, C. M., Pascale, F., Civalleri, B., Doll, K., Harrison, N. M., Bush, I. J., D'Arco, P. & Llunell, M. (2009). *CRYSTAL09 User's Manual*. University of Turin, Turin, Italy.

Dunitz, J. D. (1996). *X-ray Analysis and the Structure of Organic Molecules*. Basel, Weinheim, New York: Wiley-VCH.

Dunning, T. H. (1989). *J. Chem. Phys.* **90**, 1007–1023.

El Haouzi, A., Hansen, N. K., Le Hénaff, C. & Protas, J. (1996). *Acta Cryst.* A**52**, 291–301.

Espinosa, E., Lecomte, C. & Molins, E. (1999). *Chem. Phys. Lett.* **300**, 745–748.

Espinosa, E., Molins, E. & Lecomte, C. (1998). *Chem. Phys. Lett.* **285**, 170–173.

Farrugia, L. J. & Evans, C. (2005). *J. Phys. Chem. A*, **109**, 8834–8848.

Feast, G. C., Haestier, J., Page, L. W., Robertson, J., Thompson, A. L. & Watkin, D. J. (2009). *Acta Cryst.* C**65**, o635–o638.

Fischer, A., Tiana, D., Scherer, W., Batke, K., Eickerling, G., Svendsen, H., Bindzus, N. & Iversen, B. B. (2011). *J. Phys. Chem. A*, **115**, 13061–13071.

Frisch, M. J. *et al.* (2009). *GAUSSIAN09.* Gaussian Inc., Pittsburgh, Pennsylvania, USA.

Gavezzotti, A. (2002). *J. Phys. Chem. B*, **106**, 4145–4154.

Gavezzotti, A. (2003). *J. Phys. Chem. B*, **107**, 2344–2353.

Grabowsky, S., Kalinowski, R., Weber, M., Förster, D., Paulmann, C. & Luger, P. (2009). *Acta Cryst.* B**65**, 488–501.

Grimme, S. (2004). *J. Comput. Chem.* **25**, 1463–1473.

Grimme, S. (2006). *J. Comput. Chem.* **27**, 1787–1799.

Guillot, B. (2012). *Acta Cryst.* A**68**, s204.

Guzei, I. A., Bikzhanova, G. A., Spencer, L. C., Timofeeva, T. V., Kinnibrugh, T. L. & Campana, C. F. (2008). *Cryst. Growth Des.* **8**, 2411–2418.

Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* A**34**, 909–921.

Herbstein, F. H. (2000). *Acta Cryst.* B**56**, 547–557.

Hirshfeld, F. L. (1976). *Acta Cryst.* A**32**, 239–244.

Hirshfeld, F. L. & Rabinovich, D. (1973). *Acta Cryst.* A**29**, 510–513.

Hoser, A. A., Dominiak, P. M. & Woźniak, K. (2009). *Acta Cryst.* A**65**, 300–311.

Howard, S. T., Hursthouse, M. B., Lehmann, C. W. & Poyner, E. A. (1995). *Acta Cryst.* B**51**, 328–337.

Jarzembska, K. N. & Dominiak, P. M. (2012). *Acta Cryst.* A**68**, 139–147.

Jarzembska, K. N., Kubsik, M., Kamiński, R., Woźniak, K. & Dominiak, P. M. (2012). *Cryst. Growth Des.* **12**, 2508–2524.

Jayatilaka, D. & Grimwood, D. J. (2001). *Acta Cryst.* A**57**, 76–86.

Jelsch, C., Guillot, B., Lagoutte, A. & Lecomte, C. (2005). *J. Appl. Cryst.* **38**, 38–54.

Kamiński, R., Herbaczyńska, B., Srebro, M., Pietrzykowski, A., Michalak, A., Jerzykiewicz, L. B. & Woźniak, K. (2011). *Phys. Chem. Chem. Phys.* **13**, 10280–10284.

Kamiński, R., Jarzembska, K. N. & Domagała, S. (2013). *J. Appl. Cryst.* **46**, 540–543.

Keen, D. A., Gutmann, M. J. & Wilson, C. C. (2006). *J. Appl. Cryst.* **39**, 714–722.

Krishnan, R., Binkley, J. S., Seeger, R. & Pople, J. A. (1980). *J. Chem. Phys.* **72**, 650–654.

Kuntzinger, S., Dahaoui, S., Ghermani, N. E., Lecomte, C. & Howard, J. A. K. (1999). *Acta Cryst.* B**55**, 867–881.

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.

Madsen, A. Ø. (2006). *J. Appl. Cryst.* **39**, 757–758.

Madsen, A. & Larsen, S. (2007). *Angew. Chem. Int. Ed.* **46**, 8609–8613.

Madsen, A. Ø, Sørensen, H. O., Flensburg, C., Stewart, R. F. & Larsen, S. (2004). *Acta Cryst.* A**60**, 550–561.

Makal, A., Benedict, J., Trzop, E., Sokolow, J., Fournier, B., Chen, Y., Kalinowski, J. A., Graber, T., Henning, R. & Coppens, P. (2012). *J. Phys. Chem. A*, **116**, 3359–3365.

Makal, A., Trzop, E., Sokolow, J., Kalinowski, J., Benedict, J. & Coppens, P. (2011). *Acta Cryst.* A**67**, 319–326.

Martin, A. & Pinkerton, A. A. (1998). *Acta Cryst.* B**54**, 471–477.

Meindl, K. & Henn, J. (2008). *Acta Cryst.* A**64**, 404–418.

Meindl, K., Henn, J., Kocher, N., Leusser, D., Zachariasse, K. A., Sheldrick, G. M., Koritsanszky, T. & Stalke, D. (2009). *J. Phys. Chem. A*, **113**, 9684–9691.

Møller, C. & Plesset, M. S. (1934). *Phys. Rev.* **46**, 618–622.

Munshi, P., Madsen, A. Ø., Spackman, M. A., Larsen, S. & Destro, R. (2008). *Acta Cryst.* A**64**, 465–475.

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.

Nelyubina, Y. V., Antipin, M. Y., Cherepanov, I. A. & Lyssenko, K. A. (2010). *CrystEngComm*, **12**, 77–81.

Nelyubina, Y. V., Dalinger, I. L. & Lyssenko, K. A. (2011). *Angew. Chem. Int. Ed.* **50**, 2892–2894.

Paciorek, W. A., Meyer, M. & Chapuis, G. (1999). *Acta Cryst.* A**55**, 543–557.

Paul, A., Kubicki, M., Jelsch, C., Durand, P. & Lecomte, C. (2011). *Acta Cryst.* B**67**, 365–378.

Petříček, V., Dušek, M. & Palatinus, L. (2006). *JANA2006.* Institute of Physics, Czech Academy of Sciences, Prague, Czech Republic.

Poulain-Paul, A., Nassour, A., Jelsch, C., Guillot, B., Kubicki, M. & Lecomte, C. (2012). *Acta Cryst.* A**68**, 715–728.

Sands, D. E. (1982). *Vectors and Tensors in Crystallography.* London, Amsterdam, Toronto, Sydney, Tokyo: Addison-Wesley Publishing Company.

Schmøkel, M. S., Bjerg, L., Overgaard, J., Larsen, F. K., Madsen, G. K. H., Sugimoto, K., Takata, M. & Iversen, B. B. (2013). *Angew. Chem. Int. Ed.* **52**, 1503–1506.

Schmøkel, M. S., Cenedese, S., Overgaard, J., Jørgensen, M. R., Chen, Y. S., Gatti, C., Stalke, D. & Iversen, B. B. (2012). *Inorg. Chem.* **51**, 8607–8616.

Schmøkel, M. S., Kamiński, R., Benedict, J. B. & Coppens, P. (2010). *Acta Cryst.* A**66**, 632–636.

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

Schwarzenbach, D., Abrahams, S. C., Flack, H. D., Gonschorek, W., Hahn, Th., Huml, K., Marsh, R. E., Prince, E., Robertson, B. E., Rollett, J. S. & Wilson, A. J. C. (1989). *Acta Cryst.* A**45**, 63–75.

Seiler, P., Schweizer, W. B. & Dunitz, J. D. (1984). *Acta Cryst.* B**40**, 319–327.

Shapiro, S. S. & Wilk, M. B. (1965). *Biometrika*, **52**, 591–611.

Spackman, M. A. (1992). *Chem. Rev.* **92**, 1769–1797.

Spackman, M. A. & Byrom, P. G. (1996). *Acta Cryst.* B**52**, 1023–1035.

Spackman, M. A. & Byrom, P. G. (1997a). *Acta Cryst.* B**53**, 553–564.

Spackman, M. A. & Byrom, P. G. (1997b). *Chem. Phys. Lett.* **267**, 215–220.

Spackman, M. A. & Jayatilaka, D. (2009). *CrystEngComm*, **11**, 19–32.

Spackman, M. A., McKinnon, J. J. & Jayatilaka, D. (2008). *CrystEngComm*, **10**, 377–388.

Stash, A. & Tsirelson, V. (2002). *J. Appl. Cryst.* **35**, 371–373.

Stash, A. I. & Tsirelson, V. G. (2005). *Crystallogr. Rep.* **50**, 177–184.

Stevens, E. D. & Coppens, P. (1980). *Acta Cryst.* B**36**, 1864–1876.

Stevens, E., Coppens, P., Feld, R. & Lehmann, M. (1979). *Chem. Phys. Lett.* **67**, 541–543.

Su, Z. & Coppens, P. (1998). *Acta Cryst.* A**54**, 646–652.

Treutler, O. & Ahlrichs, R. (1995). *J. Chem. Phys.* **102**, 346.

Volkov, A., Abramov, Y. A. & Coppens, P. (2001). *Acta Cryst.* A**57**, 272–282.

Volkov, A., King, H. F. & Coppens, P. (2006). *J. Chem. Theory Comput.* **2**, 81–89.

Volkov, A., Li, X., Koritsanszky, T. & Coppens, P. (2004). *J. Phys. Chem. A*, **108**, 4283–4300.

Waser, J. (1955). *Acta Cryst.* **8**, 731.

Watkin, D. (2008). *J. Appl. Cryst.* **41**, 491–522.

Zarychta, B., Zaleski, J., Kyzioł, J., Daszkiewicz, Z. & Jelsch, C. (2011). *Acta Cryst.* B**67**, 250–262.

Zhurov, V. V., Zhurova, E. A. & Pinkerton, A. A. (2011). *Inorg. Chem.* **50**, 6330–6333.

Zhurov, V. V., Zhurova, E. A., Stash, A. I. & Pinkerton, A. A. (2011a). *Acta Cryst.* A**67**, 160–173.

Zhurov, V. V., Zhurova, E. A., Stash, A. I. & Pinkerton, A. A. (2011b). *J. Phys. Chem. A*, **115**, 13016–13023.

Radosław Kamiński *et al.* · Statistical analysis of structural parameters   **91**