research papers

Improvement of anisotropic displacement parameters from invariom-model refinements for three L-hydroxylysine structures

**B. Dittrich,[a]* J. J. McKinnon[b] and J. E. Warren[c]**

[a]Institut für Anorganische Chemie der Universität Göttingen, Tammannstrasse 4, Göttingen, D 37077, Germany, [b]Chemistry – M313, School of Biomedical, Biomolecular and Chemical Sciences, University of Western Australia, Crawley, WA 6009, Australia, and [c]SRS, Daresbury Laboratory, Warrington, Cheshire WA4 4AD, England

Correspondence e-mail: bdittri@gwdg.de

Three L-hydroxylysine structures have been determined at 100 K by single-crystal X-ray diffraction. High-resolution data using either a laboratory or synchrotron source were collected and subjected to invariom- and independent atom-model (IAM) refinements. Anisotropic displacement parameters (ADPs) obtained from invariom refinement were compared (i) with results from a full multipole and (ii) with an IAM high-order refinement. Differences were visualized with the program *PEANUT* and were complemented by quantitative results from a Hirshfeld test. Influences of scale factor differences, and of refinement against F^2 *versus* F, have been investigated. Systematic errors were observed in the IAM, especially when only low-order data were available. Although these errors were reduced in high-order IAM refinements, they only disappeared in charge density – and likewise – invariom refinements.

1. Introduction

L-Hydroxylysine [(2S,5R)-2,6-diamino-5-hydroxyhexanoic acid] is a non-standard amino acid that occurs, after hydroxyproline, as the second most common amino acid in collagen, the main protein of connective tissue in animals and the most abundant protein (up to 25%) in mammals. Biosynthesis of hydroxylysine in the body is achieved by the enzyme lysyl hydroxylase, a peripheral membrane protein which is part of the lumen of the endoplasmic reticulum.
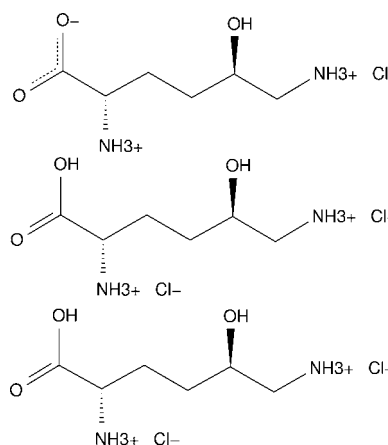
Hydroxylysine has, compared with lysine, an additional hydroxy group at Cδ (IUPAC-IUB Commission on Biochemical Nomenclature, 1970). An X-ray crystal structure of L-hydroxylysine has not been previously determined; in our attempts to crystallize it we found a hydrochloride (I), a

Table 1
Crystal data of L-hydroxylysine hydrochloride (I), L-hydroxylysine dihydrochloride (II) and L-hydroxylysine dihydrochloride monohydrate (III).

	(I)	(II)	(III)
Crystal data			
Chemical formula	$C_6H_{15}N_2O_3 \cdot Cl$	$C_6H_{16}N_2O_3 \cdot 2Cl$	$C_6H_{16}N_2O_3 \cdot 2Cl \cdot H_2O$
Cell setting, space group	Orthorhombic, $P2_12_12_1$	Monoclinic, $P2_1$	Triclinic, $P1$
Formula weight	198.65	235.11	253.12
Temperature (K)	100	100	100
a, b, c (Å)	5.4315 (1), 9.8019 (1), 17.3505 (2)	8.6224 (2), 7.0489 (2), 9.8061 (2)	5.4491 (3), 7.3962 (4), 8.1337 (4)
α, β, γ (°)	90, 90, 90	90, 106.486 (2), 90	79.382 (4), 73.191 (5), 69.544 (5)
V (Å^3)	923.72 (2)	571.50 (3)	292.75 (3)
Z	4	2	1
D_x (Mg m^{-3})	1.428	1.366	1.436
Radiation type	Synchrotron	Mo $K\alpha$	Mo $K\alpha$
μ (mm^{-1})	0.39	0.55	0.55
Crystal form, colour	Rectangle, colourless	Rectangle, colourless	Rectangle, colourless
Crystal size (mm)	$0.19 \times 0.18 \times 0.11$	$0.37 \times 0.25 \times 0.22$	$0.44 \times 0.40 \times 0.18$
Data collection			
Diffractometer	Bruker D8	Oxford Diffraction Xcalibur S	Oxford Diffraction Xcalibur S
Data collection method	ω scans	ω scans	ω scans
Absorption correction	Empirical	Analytical	Analytical
T_{min}	0.654	0.878	0.845
T_{max}	0.748	0.916	0.924
No. of measured, independent and observed reflections	232 425, 10 455, 9078	29 885, 9042, 6540	45 855, 12 708, 10 693
Criterion for observed reflections	$F > 3\sigma(F)$	$F > 3\sigma(F)$	$F > 3\sigma(F)$
R_{int}	0.096†	0.036	0.022
θ_{max} (°)	40.0	45.3	51.4
Refinement			
Refinement on	F	F	F
$R[F > 3\sigma(F)]$, $wR(F)$, S	0.031, 0.034, 1.90	0.023, 0.018, 1.73	0.020, 0.019, 1.73
No. of reflections	9078	6540	10 693
No. of parameters including H parameters	169	181	190
H-atom treatment	Constrained‡	Constrained‡	Constrained‡
Weighting scheme	$w = 1/[\sigma^2(F_o)]$	$w = 1/[\sigma^2(F_o)]$	$w = 1/[\sigma^2(F_o)]$
$(\Delta/\sigma)_{max}$	< 0.0001	< 0.0001	< 0.0001
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	1.67, −0.28	0.60, −0.25	0.70, −0.23

Computer programs used: *APEX*2 (Bruker AXS Inc., 2004), *CrysAlis CCD* and *RED* (Oxford Diffraction Ltd, 2006), *SHELXS* (Sheldrick, 2008). † To be able to measure strong low-order reflections exceeding the dynamic range of the CCD detector an Al-filter and short exposure times were required. This led to a higher R_{int} value when these data were merged with re-measured data using the same settings without the Al filter. ‡ Constrained to parent site.

dihydrochloride (II) and a dihydrochloride co-crystallizing with water (III). All three datasets were measured to relatively high resolution and were subject to invariom modelling (Dittrich *et al.*, 2004), invoking the Hansen & Coppens multipole model (Hansen & Coppens, 1978). On two of the three structures multipole refinements were performed for comparison.

The modelling of experimental diffraction data with transferable experimentally derived non-spherical scattering factors was first performed in the early nineties (Brock *et al.*, 1991). Koritsanszky *et al.* (2002) have shown that non-spherical scattering factors can be obtained by following a purely theoretical methodology, allowing the construction of theoretical aspherical-atom databases. The invariom database (Dittrich, Hübschle *et al.*, 2006), used throughout this paper, and the University of Buffalo database (Dominiak *et al.*, 2007) are such databases, whilst there is also a database based on experimental diffraction data (Pichon-Pesme *et al.*, 1995; Zarychta *et al.*, 2007). Various applications of the experimental and the theoretical databases (*e.g.* Jelsch *et al.*, 1998; Dittrich, Hübschle *et al.*, 2006; Volkov *et al.*, 2007) have been reported.

Figures-of-merit as well as anisotropic displacement parameters of the three L-hydroxylysine structures were compared with those from the independent-atom model (IAM). As pointed out earlier (Dittrich *et al.*, 2007), improvements in the figures-of-merit of a charge density or invariom refinement can be modest, *i.e.* a reduction in the R factor of $\leq 0.5\%$, when heavier elements are present in a structure. This is a well known fact in charge-density analysis (Koritsánszky *et al.*, 1994). It is due to the heavy-element core-scattering contribution to the diffraction pattern and also holds for the structures investigated here. The amount of core scattering can be approximately quantified by Stevens & Coppens' (1976) suitability factor, which therefore also gives an indication of the improvements in the figures of merit that can be expected (Dittrich *et al.*, 2007). The suitability factor indicated that with 1.60 the hydrochloride is most suited, followed by the monohydrate with 1.18. The value for the unsolvated dihydrochloride is 1.17. All three values are comparably low.

A dependence of the ADPs determined by X-ray diffraction on the scattering factor used has been observed early (Stewart, 1973). The potential problem for ADPs to be biased

Table 2
Details of invariom refinement for the three L-hydroxylysine structures.

Atom	Invariom assigned	Local symmetry	Model compound
Cl1,2	Cl⁻	6	Spherical Cl⁻
O1,3	O1c1h	m	Methanol
O2	O2c	m	Formaldehyde
N1,2	N1c1h1h1h⁺	3	Methylammonium ion
C1	C2o1o1c	m	Acetic acid
C2	C1n1c1c1h	m	2-Aminopropane
C3,4	C1c1c1h1h	mm2	Propane
C5	C1o1c1c1h	m	2-Propanol
C6	C1n1c1h1h	1	Ethylamine
H1,3	H1o[1c]	6	Methanol
H1A–C,2A–C	H1n[1c1h1h]⁺	6	Methylammonium ion
H2	H1c[1n1c1c]	6	2-Aminopropane
H3A,B,4A,B	H1c[1c1c1h]	6	Propane
H5	H1c[1o1c1c]	6	2-Propanol
H6A,B	H1c[1n1c1h]	6	Ethylamine
O1W	O1h1h	mm2	Water
H1W,2W	H1o[1h]	6	Water

by bonding density when using the IAM has been appreciated ever since ADPs were introduced (Cruickshank, 1956). Cruickshank noted that 'temperature' parameters conceal 'anisotropy of the electron density which may be due to the distribution of valence electrons'. The first study on modelling of experimental diffraction data with fixed transferable non-spherical scattering factors (Brock *et al.*, 1991) showed that the physical significance of the ADPs can be improved when bonding electron density is adequately described. A related study on a larger oligopeptide equally found a significant improvement of the thermal displacement parameters (Jelsch *et al.*, 1998). Brock & Dunitz (1982) pointed out that ADPs can either be interpreted as quantitative measures of atomic vibration or be regarded merely as convenient parameters introduced in order to improve the least-squares agreement, but not endowed with much physical significance. Refinement with non-spherical scattering factors obviates the contamination problem.

Visualizing the amount of distortion of the ADPs in the IAM was achieved by Hummel, Raselli & Bürgi (1990), who had to rely on charge-density data and a full multipole analysis for comparison. Using the invariom database (Dittrich, Hübschle *et al.*, 2006) similar and comparable results can be easily obtained in the absence of high-resolution data. By invariom modelling, contamination of ADPs due to bonding electron density is significantly reduced for standard as well as high-resolution data (Dittrich *et al.*, 2005). Visual information in the form of *PEANUT* (Hummel, Hauser & Bürgi, 1990) plots is complemented with Hirshfeld test (Hirshfeld, 1976) results.

The motivation for this study is twofold: first of all we attempt to show that invariom-model refinements yield ADPs that are equivalent to those from multipole refinements and superior to conventional high-order refinements. A second question is whether physically meaningful ADPs can also be obtained from normal datasets of limited resolution. In addition to studying ADPs we also aim to produce high-quality structural information on amino-acid molecules to generate

accurate target values for bond-distance restraints to be used in future aspherical-atom protein refinements.

2. Experimental

While single crystals of L-hydroxylysine dihydrochloride (II) were already present in the sample purchased from Sigma Aldrich, crystals of the hydrochloride (I) were grown from a solution of the compound in methanol in the absence of water. Crystals of the monohydrate (III) were grown by vapor diffusion of methanol into an aqueous saturated solution. Data collection in the home laboratory [(II) and (III)] was carried out on an Oxford Diffraction Xcalibur S diffractometer with Mo $K\alpha$ radiation, each frame covering 1° in ω. *CrysAlis Red* (Oxford Diffraction Ltd, 2006) was used for data reduction and for the face-indexed analytical absorption correction (Clark & Reid, 1995). All structures were solved with *SHELXS* (Sheldrick, 2008). Data for (I) were collected at station 9.8 at the SRS Daresbury (Cernik *et al.*, 1997). The current setup of station 9.8 consists of a Bruker APEX2 CCD detector and an open-flow Oxford Cryosystems Helix cooling device. A wavelength of 0.4750 Å was chosen to allow data collection to ultra-high resolution.

After data reduction with *SAINT* (Bruker AXS Inc., 2007) a correction for the oblique incidence (Wu *et al.*, 2002) of the X-rays with respect to the detector and an empirical absorption correction was performed with the program *SADABS*, version 05/2007 (Sheldrick, 2007). H atoms for all three structures were located in the difference-Fourier map. Crystallographic data can be found in Table 1.[1]

3. Least-squares refinements

Initial IAM refinements were performed with *SHELXL* (Sheldrick, 2008) providing starting values for subsequent aspherical-atom least-squares refinements. The atomic numbering scheme of the three molecules was chosen to be identical. Aspherical atom and IAM refinements, which included reflections with $F > 3\sigma(F)$, were subsequently performed with *XDLSM* of the *XD* package (Koritsánszky *et al.*, 2003). *XD* input files were processed with the program *InvariomTool* (Hübschle *et al.*, 2007). A full multipole refinement was not feasible for (II), as poor scattering of the crystal prevented sufficient resolution for a charge-density analysis. We therefore limited ourselves to an invariom/IAM refinement of the dataset. Scattering power was better for (III), but the low-symmetry space group $P1$ meant that data redundancy, particularly at high resolution, was low. Redundant charge-density data up to high resolution was collected for (I).

To be able to compare ADPs from different refinements we have, apart from IAM and invariom refinements, undertaken highly constrained multipole refinements of (I) and (III).

---

[1] Supplementary data for this paper are available from the IUCr electronic archives (Reference: WS5058). Services for accessing these data are described at the back of the journal.

Since the monohydrate data was of comparatively lower quality extensive use was made of local-atomic site symmetry and chemical constraints. ADPs obtained from multipole refinements were compared with those from invariom and IAM refinement.

In all refinements the chloride ions were modeled as spherical and were assigned a charge of −1. For invariom refinements non-spherical valence scattering contributions were obtained from theoretical calculations on model compounds including nearest or next-nearest neighbour atoms. Details on the model compounds used are specified in Table 2, which also contains the information on local atomic site symmetry and the chemical constraints employed in the multipole refinements. Where more than one atom is specified in the table second and subsequent atoms were constrained to the first. In the invariom refinements of (II) and (III) only eight different model compounds (excluding the Cl⁻ ion) were used to generate 14 different scattering factors – two scattering factors were added for H and O of the water molecule in the monohydrate. For (I) the atoms of the carboxylate group were modelled by the density fragments O1.5c[1.5o1c]⁻/C1.5o1.5o1c⁻ of the anion of acetic acid.

The basis set D95++(3df,3pd) was used to optimize the geometry of these model compounds with the program *GAUSSIAN*98 (Frisch *et al.*, 2002). The deviation from electroneutrality was 0.43 out of 74 valence electrons for (I), 0.83 electrons out of 82 valence electrons for (II) and 0.83 out of 90 electrons for (III). Electroneutrality was achieved by scaling monopoles of H atoms only. Full details of the general invariom modelling procedure for organic molecules can be found in Hübschle *et al.* (2007). An *ORTEP* plot of each molecular structure and the atomic labelling scheme is depicted in Fig. 1 using the result of the invariom refinement. The H-atom treatment consisted of a free refinement of positions and an isotropic displacement parameter in the IAM, whereas in the invariom and multipole refinements hydrogen-bond lengths were set to the distances obtained from geometry optimizations of the respective H-atom model compounds, only refining an isotropic displacement parameter. Optimized hydrogen-bond distances agree very well with results from neutron diffraction and can be considered equivalent.

4. Results and discussion

Fig. 1 illustrates the fact that the molecular conformation of the hydroxylysine molecule in (I) differs from (II) and (III) in the C′—Cα—Cβ—Cγ torsion angle (IUPAC-IUB Commission on Biochemical Nomenclature, 1970). Molecule (I) has a *gauche* conformation with a torsion angle of −53.7°, whereas the angle is close to 180° in (II) and (III), *i.e.* Cγ is rotated by > 60° and points towards the viewer in (I) and away from the viewer in (II) and (III). C atoms in the side chain of (I) are all *anti* and the hydroxy group is +*gauche* to Cβ.

The main conformational difference between the two dihydrochloride structures (II) and (III) is the side-chain at the hydroxy group, as characterized by the torsion angle Cβ—

Cγ—Cδ—Oε1, which is 166.9° in (II) and 58.8° in (III), meaning that Oε is almost *antiperiplanar* with respect to Cβ in (II) and *gauche* in (III), respectively.

4.1. Systematic errors in ADPs due to neglect of the aspherical density

In the following section we investigate systematic errors in ADPs that arise when the asphericity of electron density is neglected in the IAM. In the *Introduction* we mentioned that contamination of ADPs was due to bonding electron density. Cruickshank (1956) noted early on that a remedy to avoid such contamination would consist of only using high-order data in the least-squares refinement, as high-order data
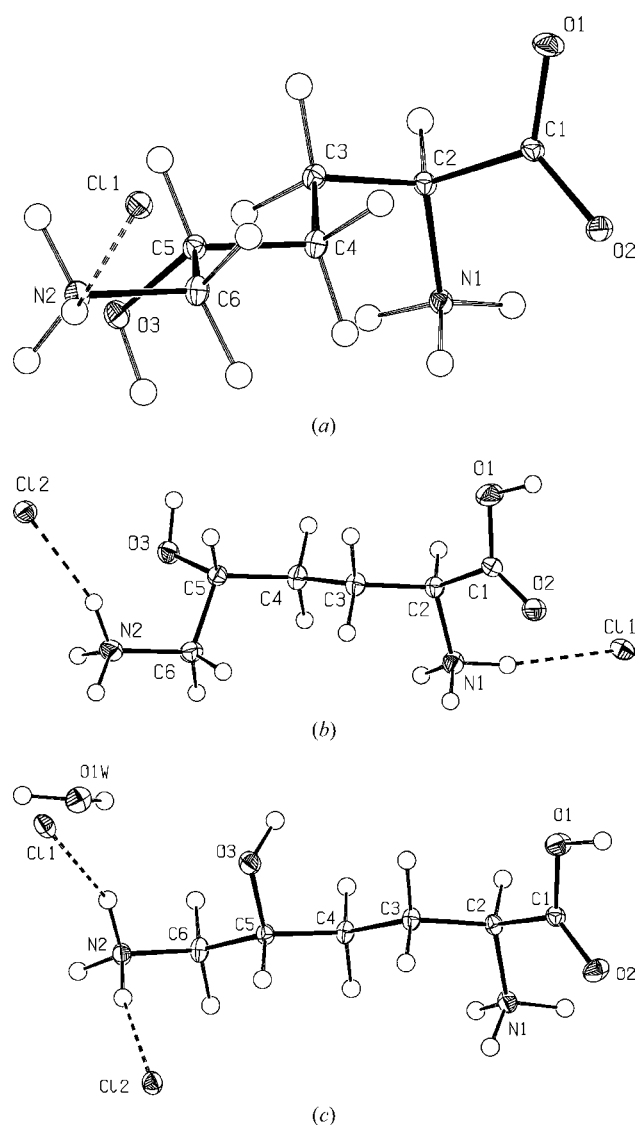

Figure 1
ORTEP representation (Burnett & Johnson, 1996) of three experimentally determined molecular structures of ʟ-hydroxylysine in the crystal with atomic numbering scheme and displacement ellipsoids at 50% probability: (*a*) ʟ-hydroxylysine hydrochloride (I), (*b*) ʟ-hydroxylysine dihydrochloride (II) and (*c*) ʟ-hydroxylysine hydrochloride monohydrate (III). Atom names were chosen to be alike in all molecules; C3 is Cβ, C4γ *etc.*, N2 is ν and O3 is Oε1.

contain more information on sharp electron density features and less information on bonding electron density.

It has been discussed previously (Hummel, Raselli & Bürgi, 1990) that differences in the ADPs are difficult to spot, making a comparison between two sets of ADPs cumbersome. To facilitate such comparisons the program *PEANUT* has been developed (Hummel, Hauser & Bürgi, 1990). A utility program was written that calculates differences between two sets of ADPs from *XD* input/result files, producing input for *PEANUT*. The utility program can be obtained from the authors of this paper. As the same crystal structures with identical unit-cell settings are investigated at the same temperature here, differences in coordinates – asphericity shifts – that were found to be relatively small between the independent atom and the pseudoatom model (Coppens *et al.*, 1969) were not taken into account.

In Fig. 2 we calculated the difference in ADPs obtained from IAM refinement and high-order IAM refinement including only reflections with a resolution in $\sin\theta/\lambda$ higher than 0.4 Å^{-1}. The relatively low cutoff was chosen to allow a comparison between all three data sets, since data for (II) are of limited resolution. As the IAM-refined ADPs are usually larger than those from either a high-order IAM or an invariom refinement, subsequent difference plots were always created by subtracting high-order IAM or invariom-model ADPs from the IAM ADPs using the IAM geometry. An overestimation of the ADPs appears in blue, whereas an underestimation is red. Fig. 2 shows that there is a considerable contamination of the ADPs in the IAM refinements for all three molecules. The amount of contamination is larger for (II) than for (I) and (III). We suspect that resolution-dependent scaling problems due to systematic errors as discussed below are the reason for
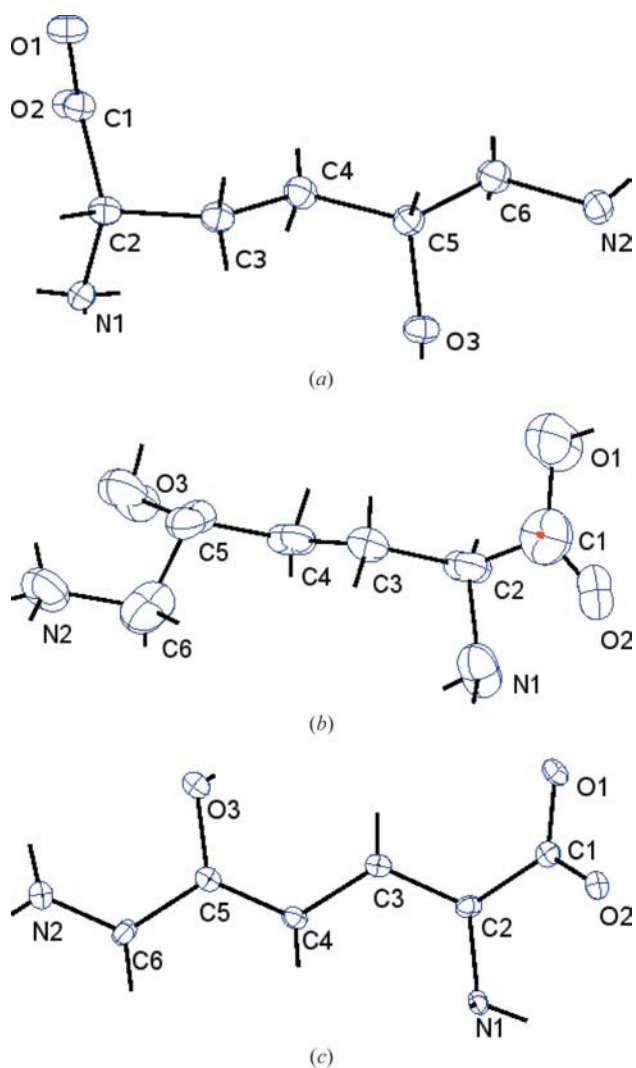

Figure 2
PEANUT representations (Hummel, Hauser & Bürgi, 1990) of the difference between ADPs obtained from an IAM refinement and a high-order IAM refinement of L-hydroxylysine hydrochloride (I), L-hydroxylysine dihydrochloride (II) and the L-hydrochloride dihydrochloride monohydrate (III). A root-mean-square difference (r.m.s.d.) scale of 10 was used. Considerable contamination of the ADPs in the IAM becomes obvious.
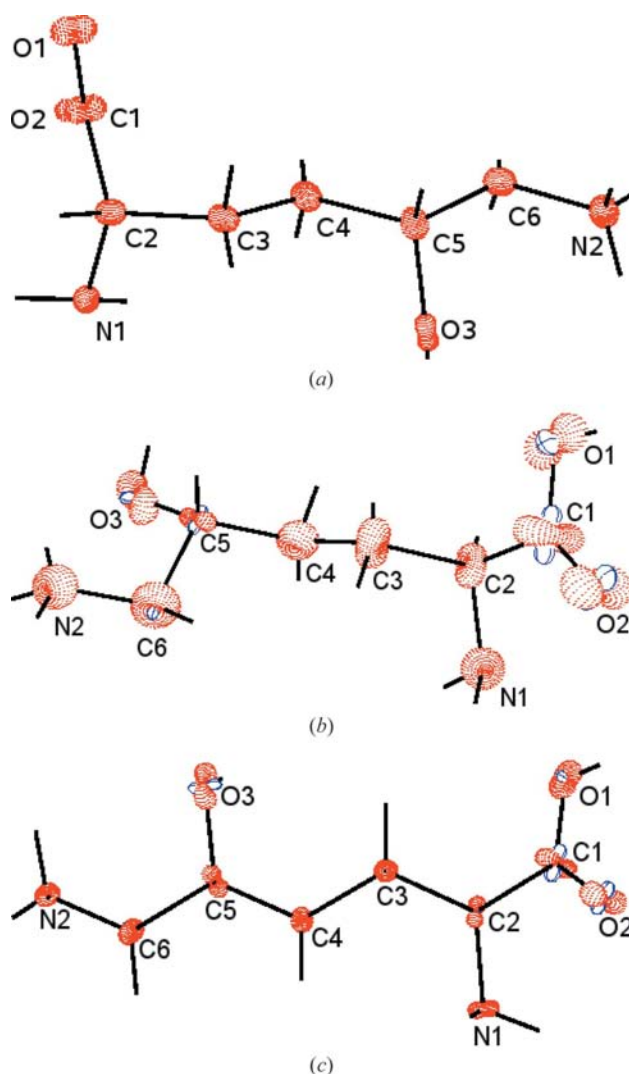

Figure 3
PEANUT representations (Hummel, Hauser & Bürgi, 1990) of the difference between ADPs obtained from a high-order IAM refinement and an invariom refinement of L-hydroxylysine hydrochloride (I), L-hydroxylysine dihydrochloride (II) and L-hydroxylysine dihydrochloride monohydrate (III). A root-mean-square difference (r.m.s.d.) scale of 10 was used. Differences are smaller than between IAM and high-order IAM refinements.

Table 3
Hirshfeld test results for L-hydroxylysine hydrochloride (I), L-hydroxylysine dihydrochloride (II) and L-hydroxylysine dihydrochloride monohydrate (III).

The difference between mean-square displacement amplitudes (DMSDA, in Å^2) is reduced in invariom (inv) and multipole (mult) refinements. Distances d (in Å) are also given.

Bond	d_{inv}	DMSDA$_{inv}$	d_{IAM}	DMSDA$_{IAM}$	d_{mult}	DMSDA$_{mult}$
(I)						
O1—C1	1.2518 (5)	2	1.2526 (6)	6	1.2510 (5)	1
O2—C1	1.2627 (5)	5	1.2634 (6)	8	1.2623 (5)	3
O3—C5	1.4269 (5)	6	1.4280 (6)	7	1.4263 (5)	5
N1—C2	1.4895 (5)	0	1.4892 (6)	0	1.4879 (5)	0
N2—C6	1.4947 (6)	−1	1.4948 (6)	−1	1.4925 (6)	0
C1—C2	1.5339 (5)	1	1.5341 (5)	1	1.5317 (5)	1
C2—C3	1.5333 (6)	0	1.5328 (6)	0	1.5320 (5)	0
C3—C4	1.5345 (6)	2	1.5340 (7)	3	1.5314 (5)	2
C4—C5	1.5340 (6)	−1	1.5333 (6)	−1	1.5319 (5)	−1
C5—C6	1.5241 (6)	5	1.5236 (7)	6	1.5214 (6)	6
Average		2.3		3.3		1.9
(II)						
O1—C1	1.3279 (9)	−20	1.3241 (13)	11	–	–
O2—C1	1.1962 (8)	22	1.2057 (10)	24	–	–
O3—C5	1.4216 (9)	−7	1.4276 (12)	−2	–	–
N1—C2	1.4872 (9)	−6	1.4829 (13)	−16	–	–
N2—C6	1.4889 (10)	−8	1.4891 (13)	−10	–	–
C1—C2	1.5155 (10)	−20	1.5124 (13)	−23	–	–
C2—C3	1.5373 (10)	−3	1.5366 (13)	−7	–	–
C3—C4	1.5284 (10)	2	1.5256 (13)	−5	–	–
C4—C5	1.5276 (10)	−7	1.5236 (14)	0	–	–
C5—C6	1.5213 (10)	2	1.5101 (14)	10	–	–
Average		9.7		10.8	–	–
(III)						
O1—C1	1.3164 (6)	−15	1.3144 (7)	5	1.3175 (7)	−2
O2—C1	1.2121 (5)	6	1.2177 (6)	6	1.2159 (7)	1
O3—C5	1.4310 (6)	−2	1.4314 (7)	6	1.4321 (7)	5
N1—C2	1.4852 (6)	3	1.4842 (7)	−5	1.4842 (7)	4
N2—C6	1.4953 (6)	−1	1.4955 (7)	−4	1.4933 (6)	1
C1—C2	1.5236 (6)	−2	1.5225 (7)	−4	1.5215 (6)	−1
C2—C3	1.5329 (6)	−5	1.5324 (7)	−8	1.5335 (6)	−7
C3—C4	1.5319 (6)	1	1.5323 (7)	1	1.5311 (6)	1
C4—C5	1.5259 (6)	1	1.5256 (7)	−1	1.5252 (6)	−3
C5—C6	1.5217 (6)	5	1.5221 (7)	6	1.5239 (6)	3
Average		4.1		4.6		2.8

IAM high-order refinement is to be avoided to prevent an electron-density bias when the multipole parameters are later refined.

The next area we want to study is the ADP differences between the high-order IAM refinement as described earlier and the invariom model that includes all the data. In Fig. 3 these differences are depicted for all molecules. It can be noticed that overall differences are smaller than for IAM minus high-order IAM (Fig. 2), but that systematic differences do likewise occur at all sites. The invariom model takes into account atomic charges. This is reflected in a systematic underestimation of the ADPs from the high-order IAM refinement as seen in a prevalence of red features. Therefore, even when IAM ADPs look spherical and physically reasonable, they might still be affected by systematic error. Systematic differences at the O atom might have further origins; it is interesting to note that the directions of the positive and negative differences are similar in (II) and (III). We think that the differences at the oxygen sites are due to the fact that invariom modelling includes the sharp features of the oxygen lone-pair density, whereas the high-order IAM refinement does not. A further possibility is that known shortcomings of the multipole model in describing diffuse electron-density distributions become visible (Spackman & Byrom, 1996; Volkov & Coppens, 2001), and this might be another contributing but probably minor factor here. Overall the invariom model ADPs can be considered to be of better quality than those from a high-order IAM refinement. This aspect is of strong interest for normal structure determinations, where high-order data are not available.

As data resolution and quality for (I) and (III) was sufficient, we have performed constrained multipole refinements to obtain benchmark ADPs for comparison. Fig. 4 shows a comparison of the difference between the ADPs from multipole and invariom refinements using the full datasets. Differences are smallest with respect to all other earlier comparisons and show that invariom ADPs are equivalent to those obtained from a full multipole refinement. Some differences in the region of the oxygen lone-pair density remain and we have discussed the possible reasons for disagreement above. Here these differences could also be a manifestation of the hydrogen-bonding environment of the respective O atoms, which is taken into account in the multipole, but not in the invariom refinement. Overall we can conclude that the invariom and multipole ADPs are equivalent when using diffraction data of the same resolution; differences are of the same order of magnitude in blue and red.

Our analysis has so far focused on the qualitative aspects of ADP differences that have been visualized in *PEANUT* plots
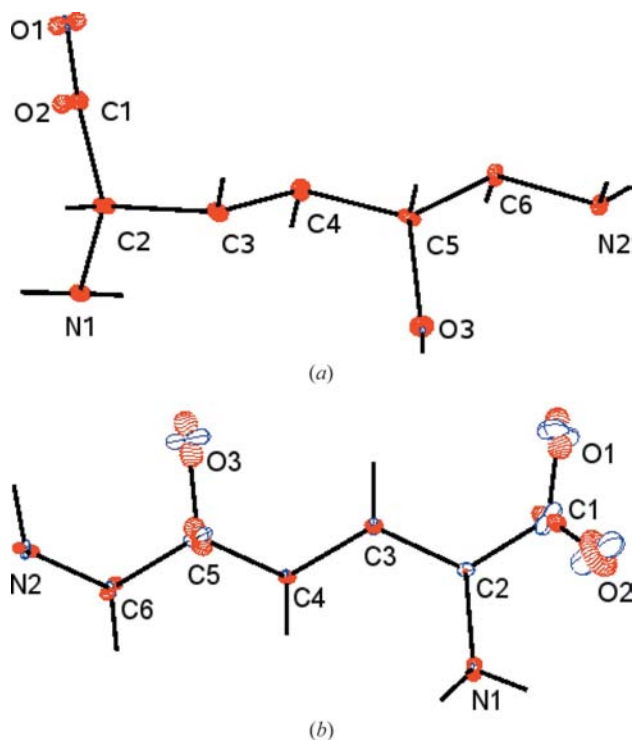
this, since the number of observed reflections in the higher-resolution shells is smaller for (II) (Table 1). It is also possible that the restricted high-order data for (II) are less accurate. A well known conclusion is that high-order data are absolutely necessary for multipole refinements. This finding supports a minimal resolution requirement of $\sin\theta/\lambda \geq 1.1$ Å in charge-density refinement (Koritsánszky *et al.*, 1998). Another point of interest is the fact that ADPs in the IAM are overestimated, since additional electron density due to charge transfer from the H atoms to C, N or O atoms is neglected. The fact that IAM ADPs model bonding electron density is visible in systematic differences often pointing in the direction of a bonding partner. Such directionality will become more obvious in subsequent *PEANUT* plots. When there are several covalent bonds of equal strength, ADP differences are less easy to understand, but can be rationalized as a vector sum of the contribution of the respective bonds. Although high-order refinements provide significantly better ADPs than the IAM, determination of ADPs in charge-density refinement by initial

Figure 4
PEANUT representations (Hummel, Hauser & Bürgi, 1990) of the difference between ADPs obtained from a full multipole refinement and an invariom refinement of (*a*) L-hydroxylysine hydrochloride (I) and (*b*) L-hydroxylysine dihydrochloride monohydrate (III). An r.m.s.d. scale of 10 was used. Differences are smallest with respect to all the other different sets and show that invariom ADPs are equivalent to those obtained from a full multipole refinement.

that have revealed features that might be overlooked when reading a table of numbers. We now focus our attention on the Hirshfeld test (Hirshfeld, 1976). The strength of this test is the ability to quantify ADP differences independent of individual bonding contributions. It is precisely these contributions that make the interpretation of *PEANUT* plots potentially difficult. Table 3 lists the results of the Hirshfeld tests, confirming earlier findings that in invariom refinements the average difference in mean-square displacement amplitudes (DMSDA) is reduced (Dittrich *et al.*, 2005). Most significant are the reductions in single bonds, whereas no improvements are seen for bonds of the carboxyl groups. One could speculate that in this case the choice of model compounds for invariom assignment was inadequate or could be improved. We have attempted to use a different model compound including next-nearest neighbors for the oxygen in the carboxyl group, but no improvements were found. We ascribe this to the individually different influences of the hydrogen-bonding environments; including these in the predicted scattering factors would be desirable and we will continue to investigate these matters.

4.2. The influence of data resolution

To investigate the influence of data resolution on the amount of contamination of the ADPs we have calculated (Figs. 5 and 6) the difference between the IAM and the

invariom model ADPs for the non-H atoms of (I), (II) and (III). ADP differences were calculated for three resolution cutoffs of 0.6, 0.8 and 1.0 Å^{-1} in $\sin\theta/\lambda$ in all cases. The same resolution cutoff was used for each pair of refinements.

Figs. 5, 6 and 7 show that there is a resolution dependence of the magnitude of the ADPs. When only low-resolution data are available the ADPs take into account bonding electron density or systematic errors most prominently, whereas when more high-angle data – which themselves are dominated by sharp features in the electron-density distribution like *e.g.* core electrons – are available, the ADPs have more physical meaning. This effect can already be seen when the data
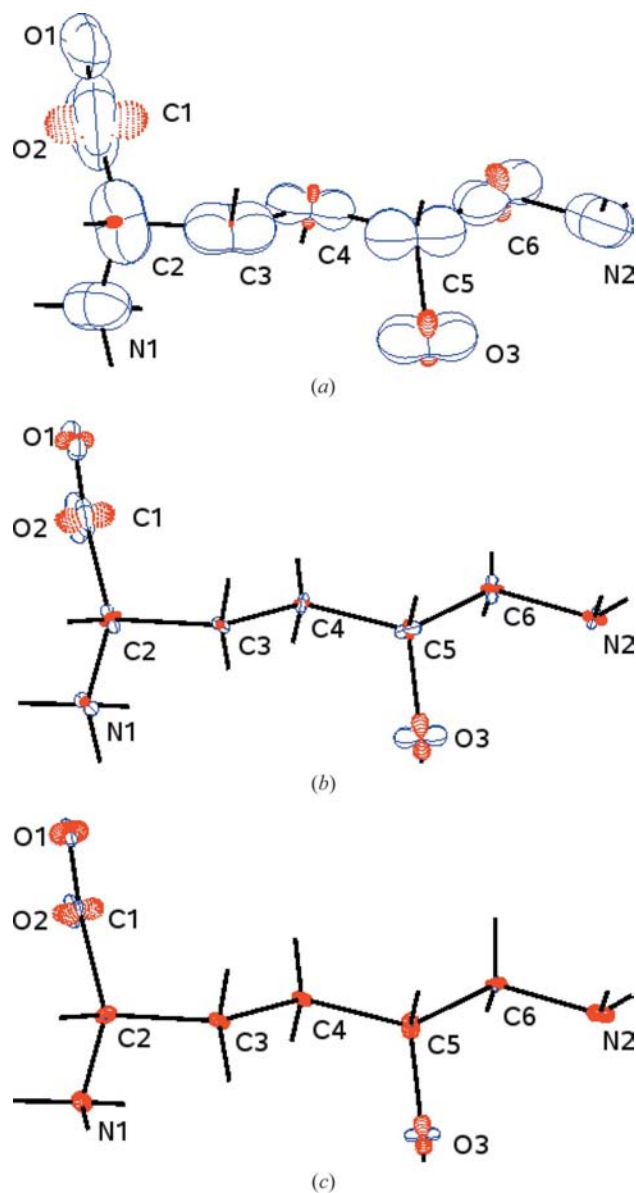

Figure 5
PEANUT representations (Hummel, Hauser & Bürgi, 1990) of the difference between the independent-atom and the invariom model of L-hydroxylysine hydrochloride (I) omitting the chloride ions: (*a*): a data cutoff of 0.6 Å^{-1} and a root-mean-square difference (r.m.s.d.) scale of 10 were used; (*b*) same r.m.s.d. but using data up to 0.8 Å^{-1}; (*c*) up to 1.0 Å^{-1}.

resolution is complete up to 0.8 Å^{-1}. However, the same systematic differences remain, if not at the same magnitude, even when the full resolution of data is used.

We believe that data resolution should not influence the magnitude of the ADPs. The fact that we can observe it for all the datasets studied here, albeit to a different extent, indicates that the data are still affected by systematic errors. An effect of the weighting scheme can be ruled out, since the weighting scheme – the same in all cases – was based on the measured uncertainties. Other sources of systematic errors have been summarized by Seiler (1992). It was pointed out that especially resolution-dependent systematic errors cause bias in ADPs (Lenstra *et al.*, 2001). Thermal diffuse scattering (TDS) might be a resolution-dependent source of error that deserves further study for area detectors, since it was possible to measure the TDS contribution to Bragg scattering with a scintillation counter (Graf *et al.*, 1981).

Volkov *et al.* (2007) mentioned that the overall scale factor can change when using aspherical scattering factors. It has been predicted that the IAM scale factor can be overestimated by up to 4% (Rees, 1978). The scale factor correlates strongly with the average magnitude of the ADPs, even more so when the data contain a resolution-dependent error. Hence, when accurate ADPs are to be refined, particular attention should be paid to avoid systematic errors. We conclude that the remaining subtle and often resolution-dependent systematic errors in diffraction data can cause ADP differences potentially exceeding the model-dependent differences discussed above. Even with non-spherical scattering factors the absolute scale of a set of ADPs is not known accurately. Scaling procedures, especially between neutron and X-ray ADPs (Blessing, 1995), might still be required in comparative studies.

It can be argued that the practise of refinement against F in charge-density studies, while not having a direct effect on the
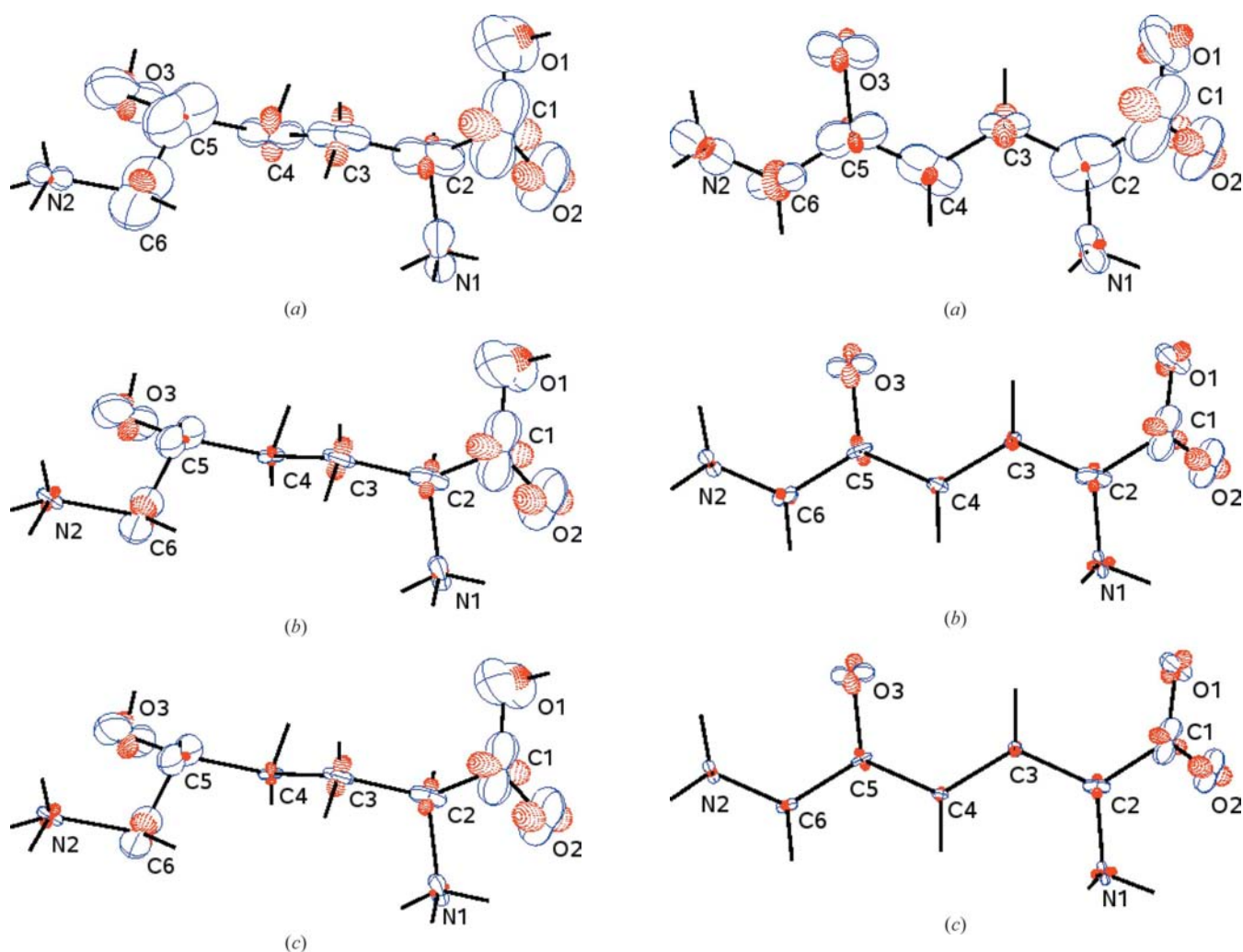

Figure 6
PEANUT representations (Hummel, Hauser & Bürgi, 1990) of the difference between the independent-atom and the invariom model of L-hydroxylysine dihydrochloride (II) omitting chloride ions: (*a*) a data cutoff of 0.6 Å^{-1} and an r.m.s.d. scale of 10 were used; (*b*) same r.m.s.d. but using data up to 0.8 Å^{-1}; (*c*) up to full resolution of 1.0 Å^{-1}.

Figure 7
PEANUT representations (Hummel, Hauser & Bürgi, 1990) of the difference between the independent-atom and the invariom model of L-hydroxylysine dihydrochloride monohydrate (III) omitting the water molecule and the chloride ions: (*a*) a data cutoff of 0.6 Å^{-1} and an r.m.s.d. scale of 10 were used; (*b*) same r.m.s.d. but using data up to 0.8 Å^{-1}; (*c*) up to 1.0 Å^{-1}.

electron density, might influence the scale factor and hence the ADPs, since in refinement against F negative intensities need to be omitted. Negative intensities might contain statistical information on measurement error. Omitting such reflections might artificially increase the scale factor that is dominated by the numerous but weak high-order data. We have compared a refinement against F to one against F^2, but did not find significant differences between them. For the determination of accurate ADPs cutting the data at a resolution level where a high I/σ ratio remains seems rather more important than refinement on F or F^2.

4.3. Absolute structure

Invariom refinement has previously been shown (Dittrich, Strümpel, Koritsánszky, Schäfer & Spackman, 2006) to reduce the standard deviations of least-squares parameters, including the Flack parameter (Flack, 1983). Due to the anomalous scattering of the chloride ions, the absolute structure and the absolute configuration of the chiral C atoms in the hydroxylysine molecules can be accurately determined from the IAM. The standard deviation of the Flack parameter, which is crucial in assessing the reliability of the analysis (Flack & Bernardinelli, 2000), was reduced from $-0.00\,(3)$ to $-0.00\,(2)$ for (II) and from $0.01\,(2)$ to $0.03\,(1)$ for (III) by invariom refinement, confirming results of the earlier study. For (I) the anomalous scattering was very small because of the short wavelength of the synchrotron radiation, so no attempt was made to determine the absolute structure.

5. Conclusion

By comparing ADPs from IAM, high-order IAM, invariom and multipole refinements for three L-hydroxylysine structures we can draw the following conclusions:

(i) As has been shown previously, ADPs from a conventional spherical-atom refinement are systematically contaminated by bonding and lone-pair electron density.

(ii) A high-order IAM refinement is not an entirely satisfactory remedy for the shortcomings of the IAM. Although systematic errors in the IAM are reduced considerably, systematic errors remain in the high-order IAM refinement. Furthermore, high-order data are not usually available.

(iii) The use of invariom allows the retrieval of additional information on thermal motion in a non-disordered crystal from standard diffraction data, although resolution-dependent systematic errors might compromise their absolute scale.

(iv) Comparison between invariom and multipole refinements shows that ADPs obtained by both methods are equivalent.

(v) Including the influence of the local hydrogen-bonding environment in the invariom scattering factors would be desirable.

When properly scaled, physically meaningful ADPs are to be reported – derived from an X-ray diffraction experiment – one should aim for the highest possible resolution maintaining a high I/σ ratio, avoiding or correcting systematic errors as

much as possible. A careful choice of specimen is imperative and non-spherical scattering factors should be used.

References

Blessing, R. H. (1995). *Acta Cryst.* B**51**, 816–823.
Brock, C. P. & Dunitz, J. D. (1982). *Acta Cryst.* B**38**, 2218–2228.
Brock, C. P., Dunitz, J. D. & Hirshfeld, F. L. (1991). *Acta Cryst.* B**47**, 789–797.
Bruker AXS Inc. (2004). *APEX*2. Bruker AXS Inc., Madison, Wisconsin, USA.
Bruker AXS Inc. (2007). *SAINT*, Version 7.45A. Bruker AXS Inc., Madison, Wisconsin, USA.
Burnett, M. N. & Johnson, C. K. (1996). *ORTEP*III. Report ORNL-6895. Oak Ridge National Laboratory, Tennessee, USA.
Cernik, R. J., Clegg, W., Catlow, C. R. A., Bushnell-Wye, G., Flaherty, J. V., Greaves, G. N., Burrows, I., Taylor, D. J., Teat, S. J. & Hamichi, M. (1997). *J. Synchrotron Rad.* **4**, 279–286.
Clark, R. C. & Reid, J. S. (1995). *Acta Cryst.* A**51**, 887–897.
Coppens, P., Sabine, T. M., Delaplane, G. & Ibers, J. A. (1969). *Acta Cryst.* B**25**, 2451–2458.
Cruickshank, D. W. J. (1956). *Acta Cryst.* **9**, 747–753.
Dittrich, B., Hübschle, C. B., Luger, P. & Spackman, M. A. (2006). *Acta Cryst.* D**62**, 1325–1335.
Dittrich, B., Hübschle, C. B., Messerschmidt, M., Kalinowski, R., Girnt, D. & Luger, P. (2005). *Acta Cryst.* A**61**, 314–320.
Dittrich, B., Koritsánszky, T. & Luger, P. (2004). *Angew. Chem. Int. Ed.* **43**, 2718–2721.
Dittrich, B., Munshi, P. & Spackman, M. A. (2007). *Acta Cryst.* B**63**, 505–509.
Dittrich, B., Strumpel, M., Schäfer, M., Spackman, M. A. & Koritsánszky, T. (2006). *Acta Cryst.* A**62**, 217–223.
Dominiak, P. M., Volkov, A., Li, X., Messerschmidt, M. & Coppens, P. (2007). *J. Chem. Theory Comput.* **2**, 232–247.
Flack, H. D. (1983). *Acta Cryst.* A**39**, 876–881.
Flack, H. D. & Bernardinelli, G. (2000). *J. Appl. Cryst.* **33**, 1143–1148.
Frisch, M. J. *et al.* (2002). *GAUSSIAN*98, Revision A.11.3. Gaussian Inc., Pittsburgh, PA, USA.
Graf, H. A., Schneider, J. R., Freund, A. K. & Lehmann, M. S. (1981). *Acta Cryst.* A**37**, 863–871.
Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* A**34**, 909–921.
Hirshfeld, F. L. (1976). *Acta Cryst.* A**32**, 239–244.
Hübschle, C. B., Luger, P. & Dittrich, B. (2007). *J. Appl. Cryst.* **40**, 623–627.
Hummel, W., Hauser, J. & Bürgi, H.-B. (1990). *J. Mol. Graphics*, **8**, 214–218.
Hummel, W., Raselli, A. & Bürgi, H.-B. (1990). *Acta Cryst.* B**46**, 683–692.
IUPAC-IUB Joint Commission on Biochemical Nomenclature (1970). *Biochemistry*, **9**, 3471–3479.
Jelsch, C., Pichon-Pesme, V., Lecomte, C. & Aubry, A. (1998). *Acta Cryst.* D**54**, 1306–1318.
Koritsánszky, T., Buschmann, J., Luger, P., Schmidt, H. & Steudel, R. (1994). *J. Phys. Chem.* **98**, 5416–5421.

Koritsánszky, T., Flaig, R., Zobel, D., Krane, H.-G., Morgenroth, W. & Luger, P. (1998). *Science*, **279**, 356–358.

Koritsánszky, T., Richter, T., Macchi, P., Volkov, A., Gatti, C., Howard, S., Mallinson, P. R., Farrugia, L., Su, Z. W. & Hansen, N. K. (2003). *XD*. Technical Report. Freie Universität Berlin, Germany.

Koritsanszky, T., Volkov, A. & Coppens, P. (2002). *Acta Cryst.* A**58**, 464–472.

Lenstra, A. T. H., Van Loock, J. F. J., Rousseau, B. & Maes, S. T. (2001). *Acta Cryst.* A**57**, 629–641.

Oxford Diffraction Ltd (2006). *CrysAlis CCD* and *CrysAlis RED*, Version 1.171.31.5. Oxford Diffraction Ltd, Abingdon, Oxfordshire, England.

Pichon-Pesme, V., Lecomte, C. & Lachekar, H. (1995). *J. Phys. Chem.* **99**, 6242–6250.

Rees, B. (1978). *Acta Cryst.* A**34**, 254–256.

Seiler, P. (1992). *Accurate Molecular Structures. Their Determination and Importance*, edited by A. Domenicano & I. Hargittai, pp. 170–198. Oxford University Press.

Sheldrick, G. M. (2007). *SADABS*, Version 2007/5. University of Göttingen, Germany.

Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.

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

Stevens, E. D. & Coppens, P. (1976). *Acta Cryst.* A**32**, 915–917.

Stewart, R. F. (1973). *Acta Cryst.* A**29**, 602–605.

Volkov, A. & Coppens, P. (2001). *Acta Cryst.* A**57**, 395–405.

Volkov, A., Messerschmidt, M. & Coppens, P. (2007). *Acta Cryst.* D**63**, 160–170.

Wu, G., Rodrigues, B. L. & Coppens, P. (2002). *J. Appl. Cryst.* **35**, 356–359.

Zarychta, B., Pichon-Pesme, V., Guillot, B., Lecomte, C. & Jelsch, C. (2007). *Acta Cryst.* A**63**, 108–125.