

Finding optimal radial-function parameters for S atoms in the Hansen–Coppens multipole model through refinement of theoretical densities

Paulina M. Dominiak* and Philip Coppens*

Department of Chemistry, State University of New York at Buffalo, New York, 14260, USA. Correspondence e-mail: pd24@buffalo.edu, coppens@buffalo.edu

Received 21 November 2005
Accepted 26 January 2006

The multipolar model of the valence-electron-density distribution for S atoms has been optimized based on theoretical structure factors for six organic molecules. The evaluation of different sets of radial function parameters shows that: (a) the ratio of the $n(l)$ values for different l is more important than their absolute values, as parallel changes in all $n(l)$ for a given atom are compensated for by a change in the refined value of the κ' parameter, and (b) the (2, 4, 6, 8) set of $n(l)$ with κ' refined as a single value is an optimal choice of radial function parameters for S atoms.

© 2006 International Union of Crystallography
Printed in Great Britain – all rights reserved

As charge-density studies are being extended to more and more complex molecules, including those containing heavier atoms, the proper choice of the density function for heavier atoms merits additional attention. This is especially true for phosphorus and sulfur which are common in molecules of biological interest.

In the Hansen–Coppens formalism (Hansen & Coppens, 1978) implemented in the *XD* program (Koritsanszky *et al.*, 2004), the pseudoatom density is defined as

$$\rho_{\kappa}(\mathbf{r}) = P_{\text{core}}\rho_{\text{core}}(r) + P_{\text{val}}\kappa^3\rho_{\text{val}}(\kappa r) + \sum_{l=1}^{l_{\text{max}}} R_l(\kappa' r) \sum_{m=1}^l P_{lm\pm} d_{lm\pm}(\mathbf{r}/r),$$

where the first term represents the spherically averaged free-atom Hartree–Fock core, ρ_{core} , and the second the spherical Hartree–Fock valence density, ρ_{val} , modified by the expansion/contraction parameter κ . The terms in the summation on the right are the atom-centered deformation functions (which may include a monopole), each described in terms of a normalized single Slater-type radial function $R_l(\kappa' r)$, multiplied by density-normalized real spherical harmonic angular functions $d_{lm\pm}(\mathbf{r}/r)$ defined in a local coordinate system. The radial functions are given by the expression

$$R_l(\kappa' r) = (\kappa'_l \alpha_l)^3 \frac{(\kappa'_l \alpha_l r)^{n(l)}}{[n(l) + 2]!} \exp(-\kappa'_l \alpha_l r),$$

where $n(l) \geq l$ to obey Poisson's electrostatic equation, and values for α_l are estimated from the Hartree–Fock optimized single- ξ exponents of the valence-orbital wavefunctions calculated for free atoms (Clementi & Raimondi, 1963; Clementi & Roetti, 1974; Macchi & Coppens, 2001).

Typically, the α_l are set to be the same for all l and are calculated as the average of the valence-orbital exponents weighted by the orbital occupations. As they are multiplied by the refinable expansion–contraction parameters κ' , the precise initial value of α_l is somewhat arbitrary.

The values of $n(l)$ may be inferred from the origin of the multipole functions as a product of atomic orbitals (Hansen & Coppens, 1978). The hydrogenic orbitals with principal quantum number n have radial

functions that are polynomials in r of order $n - 1$, with $n - l - 1$ radial nodes. In the Slater-type orbitals, the radial nodes are omitted but the power of r is preserved. As the valence-density functions for second-row atoms are products of $n = 3$ orbitals, this leads to $n(l) = 4$ for all deformation functions of such atoms. But obviously, the neglect of the radial nodes is an approximation, while the deformation functions of the pseudoatoms also must account for electron-density delocalized into the covalent bonds. It was found that, for highly accurate *Pendellösung* data on silicon and low-temperature data on NH_4SCN , a better fit is obtained when using $n(1, \dots, 4) = 4, 4, 6, 8$ (Hansen & Coppens, 1978), whereas a careful study on H_3PO_4 indicated a better fit with the (6, 6, 7, 7) set for the P atom (Moss *et al.*, 1995). The (4, 4, 6, 8) set was also used to refine experimental data for several sulfur-containing organic molecular crystals [‘hypervalent’ sulfur nitrogen species free radical (Pillet *et al.*, 2001); ‘hypervalent’ sulfur nitrogen species (Leusser *et al.*, 2004); thiapentalene (Fabius *et al.*, 1989) or tetrathiane derivatives (McCormack *et al.*, 1996)]. For thioacetamide crystals, the (4, 4, 4) set was chosen (no hexadecapoles) on the basis of a fit to theoretical structure factors (Hambley *et al.*, 2002). Also, after extensive analysis, the (4, 4, 4, 4) set was finally used in the experimental charge-density analysis of the charge-transfer complex BTDMTTF–TCNQ (Espinoza *et al.*, 1997). On the other hand, (6, 6, 7, 7) and (6, 6, 8, 8) sets, used with the KRMM method [κ -restricted multipole model (Volkov *et al.*, 2001)] were shown to give a reasonable fit to experimental data in a series of other studies (Dahaoui *et al.*, 1999; Hibbs *et al.*, 2003; Lee *et al.*, 2004; Overgaard & Hibbs, 2004).

An alternative to testing the various choices on experimental data is the evaluation of the $n(l)$ set based on electron densities as predicted by theory. Following this approach, we have generated structure factors from a series of theoretical densities of a number of isolated molecules, listed in Table 1. The *ab initio* calculations were performed at the DFT or MP2 level with *GAUSSIAN03* (Frisch *et al.*, 2004) and based on molecular geometries retrieved from the Cambridge Structure Database (CSD; Allen, 2002). All X–H bond lengths were set to average values obtained from neutron diffraction

Table 1
List of molecules included in the analysis.

Chemical diagram	Name	CSD code	Reference
	1,3-Propanedithiol	QEKXUZ	Thalladi <i>et al.</i> (2000)
	Ethyl methyl sulfide	LIQWEN	Yokoyama & Ohashi (1999)
	1-Methylthio-2-methylsulfonylethane	PINPA01	Hauback <i>et al.</i> (1994)
	<i>N</i> -Acetyl-L-cysteine	NALCYS02	Takusagawa <i>et al.</i> (1981)
	L-Cysteine	LCYSTN04	Gorbitz & Dalhus (1996)
	L-2-Amino-4,5-dithiooctanedioic acid	CUVFOO	Rajeswaran & Parthasarathy (1985)

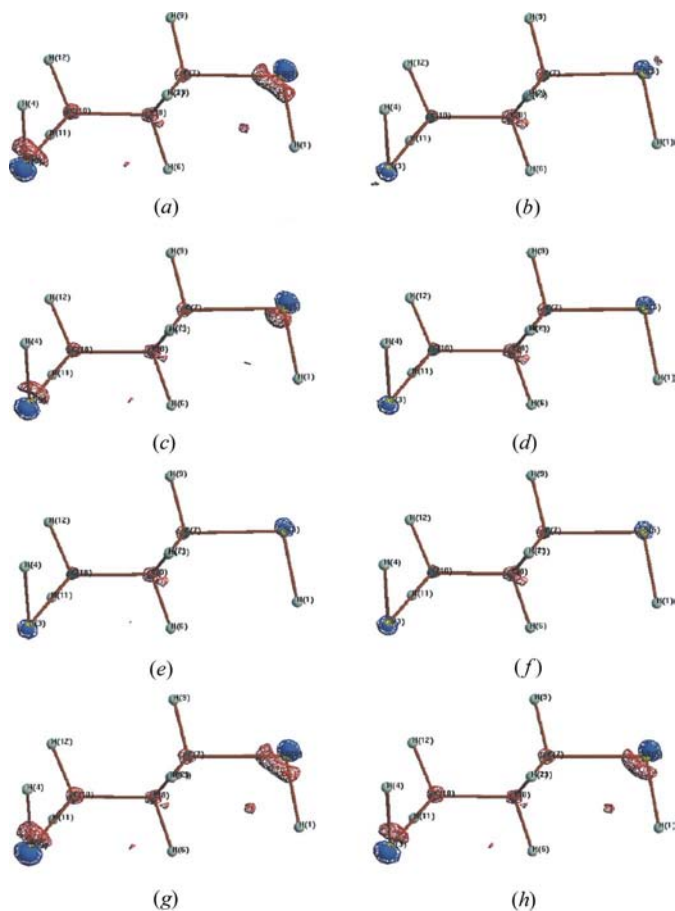


Figure 1
Three-dimensional residual electron-density maps after multipolar refinement of theoretical (B3LYP/6-31G**) valence-only structure-factor data for 1,3-propanedithiol with different sets of $n(l)$ and κ'_l values used: (a) (4, 4, 4, 4) and κ'_l the same for all l ; (b) (4, 4, 4, 4) and κ'_l separate for each l ; (c) (4, 4, 6, 8) and κ'_l the same; (d) (4, 4, 6, 8) and κ'_l separate; (e) (2, 4, 6, 8) and κ'_l the same; (f) (2, 4, 6, 8) and κ'_l separate; (g) (6, 6, 6, 6) and κ'_l the same; (h) (6, 6, 7, 7) and κ'_l the same. Contours are at the 0.1, 0.05, -0.05, -0.1, e Å⁻³ levels, blue surfaces are positive and red are negative.

(*International Tables for X-ray Crystallography*, 1992). The valence-only structure factors were obtained in the $0 < \sin \theta/\lambda < 1.1 \text{ \AA}^{-1}$ range with a metrically cubic cell of 30 Å edges. Both expansion-contraction parameters (κ , κ') were refined independently for each atom with the exception of the chemically equivalent H atoms which shared the same κ and κ' parameters. The pseudoatom expansion was truncated at the hexadecapolar level ($l_{\max} \leq 4$) for the non-H atoms and at the quadrupolar level ($l_{\max} \leq 2$) for H atoms for which only bond-directed functions of $l, m = 1, 0$ and $2, 0$ were refined. The number of independent multipole populations was restricted by imposing local symmetries. The theoretical phases were kept fixed during the refinement.

Several combinations of $n(l)$ radial-function parameters for sulfur were tested together with different basis set functions and levels of theory. Besides the standard (4, 4, 4, 4) set, we tested (4, 4, 6, 8), (6, 6, 6, 6), (6, 6, 8, 8), (6, 6, 7, 7), (5, 5, 5, 5), (7, 7, 7, 7), (8, 8, 8, 8) and (2, 4, 6, 8), the last four never examined before. Additionally, non-constrained refinement of κ'_l parameters for S atoms were performed, in which the κ' parameters were allowed to be different for different l . The α_l parameter was kept constant at $3.85125 \text{ bohr}^{-1}$.

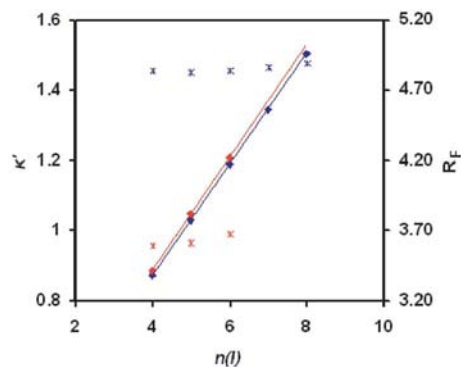


Figure 2
Correlation plot between $n(l)$, set to be equal for all l , and κ' (diamonds) and the agreement factors R_F (stars) for 1,3-propanedithiol (QEKXUZ, blue) and ethyl methyl sulfide (LIQWEN, red) molecules [QEKXUZ: $\kappa' = 0.158n(l) + 0.2411$, $R^2 = 1$; LIQWEN: $\kappa' = 0.1605n(l) + 0.2452$, $R^2 = 1$].

Table 2

Selected details for valence-only structure-factor data of six organic molecules (QEKKUZ, LIQWEN, PINPA01, NALCYS02, LCYSTN04 and CUVFOO) calculated at different levels of theory and refined with different sets of radial function parameters [$n(l)$ and κ'_l].

	$n(l)$	κ'_l	R_F	$\Delta\rho_{\min}$ $\Delta\rho_{\max}$ [e Å ⁻³]	$\kappa'_1 - \kappa'_4$
QEKKUZ					
B3LYP/6-31G**	4444	-†	4.84	-0.10 0.23	0.87‡§
	4444	+¶	3.85	-0.08 0.15	1.84 0.97 0.77 0.68
	4468	-	4.24	-0.09 0.19	1.07
	4468	+	3.77	-0.08 0.15	1.87 0.97 1.06 1.22
	5555	-	4.83	-0.10 0.22	1.03‡
	6666	-	4.84	-0.10 0.22	1.19‡
	6677	-	4.40	-0.10 0.19	1.27
	6688	-	4.23	-0.10 0.18	1.36
	7777	-	4.86	-0.10 0.22	1.35‡
	8888	-	4.89	-0.11 0.21	1.51‡
	2468	-	3.80	-0.07 0.16	1.05
	2468	+	3.68	-0.07 0.15	1.21 0.96 1.06 1.22
B3LYP/6-311++G**	4444	-	5.46	-0.11 0.31	0.85‡
	4444	+	4.54	-0.06 0.25	1.98 1.01 0.77 0.68
	4468	-	4.96	-0.09 0.30	1.09‡
	2468	-	4.49	-0.06 0.26	1.07
	2468	+	4.38	-0.05 0.24	1.31 1.00 1.06 1.23
MP2/6-311++G**	4444	-	5.12	-0.17 0.09	0.90
	4468	-	4.78	-0.16 0.07	1.09
	2468	-	4.45	-0.14 0.04	1.06
LIQWEN					
B3LYP/6-31G**	4444	-	3.59	-0.08 0.09	0.89
	4444	+	2.96	-0.08 0.06	1.45 0.99 0.74 0.66
	5555	-	3.61	-0.08 0.09	1.05
	6666	-	3.67	-0.08 0.09	1.21
	4468	-	3.18	-0.08 0.07	1.06
	2468	-	3.01	-0.08 0.06	1.03
	2468	+	2.91	-0.08 0.06	0.98 0.99 1.03 1.18
B3LYP/6-311++G**	4444	-	3.92	-0.09 0.14	0.88
	4444	+	3.27	-0.05 0.13	1.54 1.01 0.75 0.67
	4468	-	3.43	-0.05 0.13	1.07
	2468	-	3.27	-0.05 0.13	1.05
	2468	+	3.19	-0.05 0.13	1.06 1.02 1.03 1.20
PINPA01					
B3LYP/6-31G**	4444	-	3.05	-0.11 0.11	0.89 / 0.72††
	4444	+	2.68	-0.09 0.08	1.51 0.99 0.74 0.66 / 1.91 0.71 0.73 0.65
	4468	-	2.74	-0.08 0.09	1.05 / 1.03
	2468	-	2.63	-0.08 0.08	1.02 / 1.03
	2468	+	2.44	-0.07 0.08	1.02 0.99 1.04 1.20 / 1.41 0.71 1.02 1.18
B3LYP/6-311++G**	4444	-	3.19	-0.11 0.18	0.87 / 0.71
	4444	+	2.81	-0.08 0.15	1.63 1.02 0.75 0.68 / 1.65 0.55 0.72 0.65
	4468	-	2.96	-0.08 0.17	1.06 / 1.03
	2468	-	2.87	-0.08 0.16	1.04 / 1.03
	2468	+	2.65	-0.06 0.15	1.13 1.02 1.05 1.22 / 1.17 0.56 1.01 1.18
NALCYS02					
B3LYP/6-31G**	4444	-	3.24	-0.09 0.11	0.87‡
	4444	+	3.04	-0.08 0.08	1.61 0.98 0.76 0.68
	4468	-	3.13	-0.08 0.09	1.07
	2468	-	3.11	-0.08 0.08	1.04
	2468	+	3.02	-0.08 0.08	1.10 0.97 1.05 1.22
B3LYP/6-311++G**	4444	-	3.33	-0.10 0.18	0.86‡
	4444	+	3.07	-0.05 0.15	1.70 1.01 0.77 0.70
	4468	-	3.15	-0.06 0.15	1.10
	2468	-	3.07	-0.05 0.15	1.07
	2468	+	3.04	-0.05 0.14	1.19 1.01 1.06 1.26
LCYSTN04 mol1					
B3LYP/6-31G**	4444	-	3.18	-0.08 0.13	0.91‡
	4444	+	2.83	-0.09 0.09	1.54 1.00 0.75 0.68
	4468	-	2.92	-0.09 0.10	1.08
	2468	-	2.80	-0.09 0.08	1.05
	2468	+	2.76	-0.09 0.08	1.10 1.00 1.05 1.22
B3LYP/6-311++G**	4444	-	3.50	-0.10 0.18	0.90‡
	4444	+	3.13	-0.06 0.14	1.76 1.02 0.78 0.69
	4468	-	3.23	-0.06 0.16	1.07
	2468	-	3.11	-0.06 0.14	1.08
	2468	+	3.08	-0.06 0.14	1.25 1.02 1.07 1.24
LCYSTN04 mol2					
B3LYP/6-31G**	4444	-	3.24	-0.08 0.10	0.90‡
	4444	+	2.86	-0.08 0.10	1.55 0.98 0.76 0.68
	4468	-	2.98	-0.08 0.10	1.07‡
	2468	-	2.85	-0.08 0.09	1.04

Table 2 (continued)

	$n(l)$	κ'_l	R_F	$\Delta\rho_{\min}$ $\Delta\rho_{\max}$ [e Å ⁻³]	$\kappa'_1 - \kappa'_4$
B3LYP/6-311++G**	2468	+	2.81	-0.08 0.09	1.10 0.97 1.06 1.23
	4444	-	3.51	-0.10 0.18	0.89‡
	4444	+	3.18	-0.07 0.15	1.82 0.99 0.78 0.70
	4468	-	3.35	-0.08 0.17	1.10‡
	2468	-	3.18	-0.07 0.16	1.07
	2468	+	3.14	-0.07 0.15	1.29 0.99 1.08 1.26
CUVFOO					
B3LYP/6-31G**	4444	-	3.20	-0.11 0.19	0.87 / 0.88‡‡
	4444	+	2.99	-0.11 0.16	1.47 0.93 0.73 0.64 / 1.54 0.94 0.77 0.67
	4468	-	3.07	-0.11 0.19	1.02 / 1.04
	2468	-	2.97	-0.11 0.13	0.98 / 1.01
	2468	+	2.92	-0.11 0.13	0.98 0.94 1.01 1.13 / 1.03 0.95 1.09 1.19
B3LYP/6-311++G**	4444	-	3.42	-0.10 0.24	0.88 / 0.87
	4444	+	3.22	-0.08 0.21	1.53 0.94 0.75 0.64 / 1.63 0.95 0.77 0.66
	4468	-	3.30	-0.08 0.24	1.04 / 1.04
	2468	-	3.20	-0.08 0.19	1.00 / 1.02
	2468	+	3.17	-0.08 0.18	1.03 0.94 1.04 1.15 / 1.09 0.95 1.07 1.19

† – means the same for all l . ¶ + means separate refinement for each l . ‡ We were not able to refine κ' for the H atom bonded to this S atom, its κ' value was set to 1.2. § κ' is an average of two values for two thiol S atoms present in the molecule, which are essentially equal. †† Number(s) before the slash correspond(s) to the thioether S atom, after to the sulfonyl. ‡‡ Two sets of numbers correspond to two S atoms in an S bridge.

The results are summarized in Table 2. Examination of the R factors in the table shows that the best fit is obtained with the (2, 4, 6, 8) set. It may be noted that the crystallographic R value calculated for the valence-only structure factors provides a more sensitive figure of merit of the fit than the R factor for the total structure (Koritsanszky *et al.*, 2002). The better fit of the (2, 4, 6, 8) set is confirmed by the residual density maps shown in Fig. 1, which exhibit smaller residual features for this set. This result is quite independent of the basis set or level of theory applied. Only in two cases (LIQWEN and NALCYS02, B3LYP 6-311++G**) does the (4, 4, 6, 8) $\Delta\rho_{\max}$ value equal that for (2, 4, 6, 8). However, the R factor is worse and examination of the details in the residual maps shows the features to be unrealistically sharp in the former case.

The features in the residual density tend to be somewhat higher (up to ~ 0.2 e Å⁻³) for the large basis sets, which better represent sharp features in the density near the nuclei, which are not fully accommodated by the Slater-type radial functions of the multipole model. Such features are not visible in experimental data because of thermal motion.

It is quite interesting that the (4, 4, 4, 4), (5, 5, 5, 5), (6, 6, 6, 6), (7, 7, 7, 7) and (8, 8, 8, 8) sets give almost identical statistics (Table 2) and residual maps (Figs. 1a and 1g) but with different refined κ' values (Table 2). As $n(l)$ increases, *i.e.* the function becomes more diffuse, the increase in κ' , which represents a contraction, compensates for the increase in $n(l)$. This is well demonstrated by the correlation plot for κ' and $n(l)$ shown in Fig. 2. Remarkably, as a result, the R factor is practically independent of the values of $n(l)$ (Fig. 2). Clearly, the most important choice is not the $n(l)$ values, but rather their ratio $n(1) : n(2) : n(3) : n(4)$, with the higher functions being more diffuse. The higher functions are products of higher l orbitals, which have fewer nodes but are also more likely to represent bonding density.

For the (2, 4, 6, 8) model, the refinement with separate κ'_l values does not give a significant improvement over the single κ' refinement in terms of R factors and residual maps. The constrained refinement of κ' parameters with experimental data is notoriously difficult and separate refinement almost impossible. We conclude that the

(2, 4, 6, 8) set of $n(l)$ with κ' refined as a single value is an optimal choice of radial function parameters for S atoms. According to the results, this conclusion applies equally to the different types of S atoms studied, which include thiols (QEKXUZ, NALCYSO2, LCYSTN04), thioethers (LIQWEN, PINPA01), an S bridge (CUVFOO) and a sulfonyl group (PINPA01).

Support of this work by the National Science Foundation (CHE0236317) is gratefully acknowledged.

References

- Allen, F. H. (2002). *Acta Cryst.* **B58**, 380–388.
- Clementi, E. & Raimondi, D. L. (1963). *J. Chem. Phys.* **38**, 2686–2692.
- Clementi, E. & Roetti, C. (1974). *At. Data Nucl. Data Tables*, **14**, 177–478.
- Dahaoui, S., Pichon-Pesme, V., Howard, J. A. K. & Lecomte C. (1999). *J. Phys. Chem. A*, **103**, 6240–6250.
- Espinosa, E., Molins, E. & Lecomte, C. (1997). *Phys. Rev.* **56**, 1820–1833.
- Fabius, B., Cohen-Addad, C., Larsen, F. K., Lehmann, M. S. & Becker, P. (1989). *J. Am. Chem. Soc.* **111**, 5728–5732.
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J. A. Jr, Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B. *et al.* (2004). *Gaussian03* (Revision C.02). Gaussian, Inc., Wallingford, CT, USA.
- Gorbitz, C. H. & Dalhus, B. (1996) *Acta Cryst.* **C52**, 1756–1759.
- Hambley, T. W., Hibbs, D. E., Turner, P., Howard, S. T. & Hursthouse, M. B. (2002). *J. Chem. Soc. Perkin Trans. 2*, pp. 235–239.
- Hansen, N. K. & Coppens, P. (1978). *Acta Cryst.* **A34**, 909–921.
- Hauback, B. C., Brobak, T. J. & Mo, F. (1994). *Z. Kristallogr.* **209**, 60–63.
- Hibbs, D. E., Austin-Woods, C. J., Platts, J. A., Overgaard, J. & Turner P. (2003). *Chem. Eur. J.* **9**, 1075–1084.
- International Tables for Crystallography* (1992). Vol. C. Dordrecht: Kluwer Academic Publishers.
- Koritsanszky, T. S., Howard, S., Macchi, P., Gatti, C., Farrugia, L. J., Mallinson, P. R., Volkov, A., Su, Z., Richter, T. & Hansen, N. K. (2004). *XD* (version 4.12, October 2004), a Computer Program Package for Multipole Refinement & Analysis of Electron Densities from Diffraction Data. Free University of Berlin, Germany; University of Wales, Cardiff, UK; Università di Milano, Italy; CNR-ISTM, Milano, Italy; University of Glasgow, UK; State University of New York, Buffalo, USA; University of Nancy, France.
- Koritsanszky, T. S., Volkov, A. & Coppens, P. (2002). *Acta Cryst.* **A58**, 464–472.
- Lee, C.-R., Tang, T.-H., Chen, L., Wang, C.-C. & Wang, Y. (2004). *J. Phys. Chem. Solids*, **65**, 1957–1966.
- Leusser, D., Henn, J., Kocher, N., Engels, B. & Stalke, D. (2004). *J. Am. Chem. Soc.* **126**, 1781–1793.
- Macchi, P. & Coppens, P. (2001). *Acta Cryst.* **A57**, 656–662.
- McCormack, K. L., Mallinson, P. R., Webster, B. C. & Yufit, D. S. (1996). *J. Chem. Soc. Faraday Trans.* **92**, 1709–1716.
- Moss, G. R., Souhassou, M., Blessing, R. H., Espinosa, E. & Lecomte, C. (1995). *Acta Cryst.* **B51**, 650–660.
- Overgaard, J. & Hibbs, D. E. (2004). *Acta Cryst.* **A60**, 480–487.
- Pillet, S., Souhassou, M., Pontillon, Y., Caneschi, A., Gatteschi, D. & Lecomte, C. (2001). *New J. Chem.* **25**, 131–143.
- Rajeswaran, M. & Parthasarathy, R. (1985). *Acta Cryst.* **C41**, 726–728.
- Takusagawa, F., Koetzle, T. F., Kou, W. W. H. & Parthasarathy, R. (1981). *Acta Cryst.* **B37**, 1591–1596.
- Thalladi, V. R., Boese, R. & Weiss, H.-C. (2000). *J. Am. Chem. Soc.* **122**, 1186–1190.
- Volkov, A., Abramov, Y. & Coppens, P. (2001). *Acta Cryst.* **A57**, 272–282.
- Yokoyama, Y. & Ohashi, Y. (1999). *Bull. Chem. Soc. Jpn.* **72**, 2183–2191.