Acta Crystallographica Section B
Structural
Science
ISSN 0108-7681

Extracting charge density distributions from diffraction data: a model study on urea

R. Y. de Vries,[a] D. Feil[a]* and V. G. Tsirelson[b]

[a]Chemical Physics Laboratory, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands, and [b]Mendeleev University of Chemical Technology, Moscow, Russia

Correspondence e-mail: d.feil@ct.utwente.nl

Received 1 May 1999
Accepted 12 July 1999

The quality of the extraction of electron density distributions by means of a multipole refinement method is investigated. Structure factors of the urea crystal have been obtained from an electron density distribution (EDD) resulting from a density function calculation with the *CRYSTAL*95 package. To account for the thermal motion of the atoms, the stockholder-partioned densities of the atoms have been convoluted with thermal smearing functions, which were obtained from a neutron diffraction experiment. A *POP* multipole refinement yielded a good fit, $R = 0.6\%$. This disagreement factor is based on magnitudes only. Comparison with the original structure factors gave a disagreement of 0.8% owing to differences in magnitude and phase. The fitted EDD still showed all the characteristics of the interaction density. After random errors corresponding to the experimental situation were added to the structure factors, the refinement was repeated. The fit was $R = 1.1\%$. This time the resulting interaction density was heavily deformed. Repetition with another set of random errors from the same distribution yielded a widely different interaction density distribution. The conclusion is that interaction densities cannot be obtained from X-ray diffraction data on non-centrosymmetric crystals.

1. Introduction

In the past few decades it has become conventional to extract the electron density distribution (EDD) from X-ray diffraction experiments by fitting the data to a model to remove noise from the data, to overcome the incompleteness of the data set and to calculate static densities by deconvolution from the thermal motion. Various programs exist to model the EDD, *e.g. LSEXP* (Hirshfeld, 1971, 1977), *POP* (Craven *et al.*, 1987), *MOLLY* (Hansen & Coppens, 1978), *VALRAY* (Stewart & Spackman, 1983) and the relatively new *XD* (Koritsansky *et al.*, 1994) model. Assuming that neutron data will provide positional and displacement parameters, a judicious choice of multipole functions leads to a limited set of (occupation) parameters from which we can obtain a reliable deformation density. A perfect fit cannot be expected, but discrepancies will be acceptable when they are small compared with errors arising from the noise on the structure factors. Quantum chemical methods to calculate the EDD have been very useful in the development of this deformational part.

Since the introduction of these so-called multipole models there have been many studies where the EDD has been extracted from experimental data. Only very few studies were carried out to *test* these multipole models. It is interesting to see to what extent the multipole model can retrieve the EDD from a set of data calculated from a known EDD, especially in the case of non-centrosymmetric cases where the phases of the structure factors are not generally known accurately. Since the model is only fitted to the absolute value of the structure factors, there is some freedom in the phase of the structure factors in non-centrosymmetric crystals which might have an effect on the EDD that is obtained. Having a known EDD and its corresponding structure factors gives us the possibility of testing the multipole procedure and to assess the detectability of small features.

Whether the retrieved EDD has the quality necessary to draw conclusions of physical and chemical interest still has to be ascertained. We suggest using the interaction density as a criterium for the success of the retrieval procedure. This choice is not arbitrary. It reveals changes in the EDD when molecules interact; as such it is the difference between two large quantities, which is highly susceptible to errors in either one.

This interaction density plays an important role when several molecular systems interact. Thus, it is well known that the capability of a water molecule to act as an acceptor in hydrogen bonds to other molecules is increased when it is a donor at the same time (Hermansson, 1984). This phenomenon can easily be explained in terms of the change of the EDD owing to dimeric hydrogen bonding: the interaction density leads to increased dipole moments of the monomers.

Unfortunately, the interaction density distribution is not observable. As stated above, it is the difference between the EDD of a system consisting of many molecules and the superimposed EDDs of the constituting molecules. For a crystal, the EDD of the system can in principle be obtained by X-ray diffraction, but this method does not allow the determination of the EDD of isolated molecules. The computation of the interaction density also meets considerable problems. The present computing facilities make the accurate quantum chemical calculation of the EDD of small molecules with various methods possible. In the case of interacting molecules the systems rapidly become too large to allow calculations of the required accuracy. Thus, we are confronted with the problem of having a quantity that is highly relevant in the study of *e.g.* supramolecules, *i.e.* the interaction density, which is not accessible by experiment and barely by calculation. Therefore, we only have recourse to a procedure in which we subtract calculated molecular EDDs from an observed crystal EDD. Of course, puritans who claim that only the total EDD has physical significance are correct. Subtracting the molecular EDD is only useful as far as it increases insight.

We have chosen the urea crystal in our study for several reasons:

(i) urea crystallizes in a non-centrosymmetric space group, leading to phase uncertainties;

(ii) with five atoms in the asymmetric unit, the system is sufficiently small to be handled quantum mechanically with a large basis set;

(iii) hydrogen bonding with its (subtle) pronounced electrostatic features plays an important role in the crystal structure;

(iv) accurate experimental data on this crystal are available, which allows us to simulate the experimental situation closely.

We have used the *CRYSTAL*95 (Dovesi *et al.*, 1996) program to calculate the EDD and structure factors of the urea crystal.[1] The *CRYSTAL*95 program gives us the opportunity to calculate the EDD with Hartree–Fock (HF) or Density Functional Theory (DFT). We have performed both to see whether the EDDs obtained differ significantly and compared the results with the outcome of the experiment (Zavodnik *et al.*, 1999).

Recent studies by Spackman and co-workers (Spackman & Byrom, 1997; Spackman *et al.*, 1999) show that in the absence of noise on the data it is possible to retrieve the EDD of non-centrosymmetric crystals from X-ray diffraction data. Our goal is to verify the capability of the multipole method to yield the EDD at the level of the interaction density. We have tried to simulate the experimental situation as closely as possible and have to conclude that in general it is not possible to extract details at the interaction level.

2. Methods

To see to what extent the theoretically obtained EDD can be retrieved from the diffraction data with a multipole model we have taken the experimental data of Zavodnik *et al.* (1999) as the reference in our model study. This means that we have used only those structure factors (412 in number) in the refinement of the theoretical data that were reported in the paper. Furthermore, this gave us the possibility of using the experimental weights.

The HF or DFT calculation with the *CRYSTAL*95 program gives us a set of structure factors that were calculated from a static EDD. Lattice vibrations of the crystal, which consist of zero-point vibrations and thermal excitations, are not taken into account by the *CRYSTAL*95 software. In principle, it is possible to calculate the EDD at a finite temperature by calculating the wavefunction for all excited states and to make a Boltzmann distribution over all states, but in practice this is not feasible. We have approximated the thermally averaged distribution by applying the Debye–Waller factor, obtained by neutron diffraction by Swaminathan *et al.* (1984), to the atomic density distributions, obtained by partitioning the theoretical density distribution according to Hirshfeld's stockholder principle (Hirshfeld, 1977). This procedure is coded in the program *FITTER* developed by Bruning & Feil (1992).

We have chosen the *POP* multipole model to carry out the refinements. The atomic scattering curves for the Hartree–Fock *K*-shell of C, N and O were taken from Cromer & Waber (1974). Radial scattering factors for the spherical component of the C, N and O Hartree–Fock *L*-shell charge densities were constructed from a linear combination of Slater-type wavefunctions, as given by Clementi & Roetti (1974). For C, N and

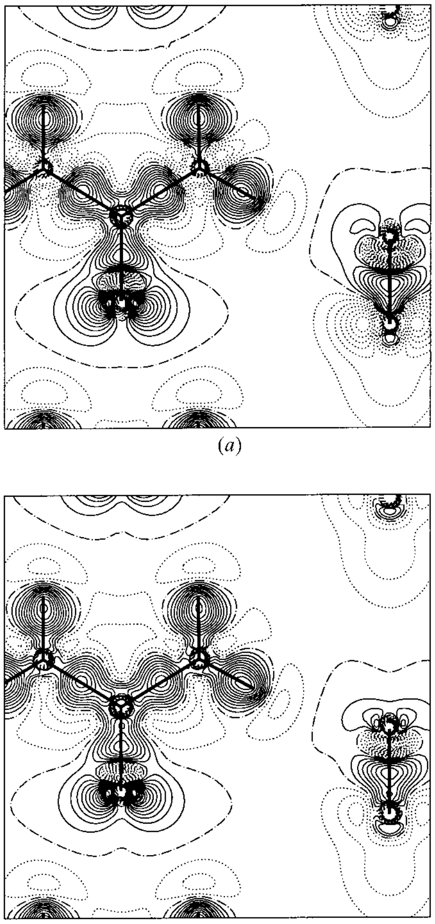
---

[1] Supplementary data for this paper are available from the IUCr electronic archives (Reference: SH0119). Services for accessing these data are described at the back of the journal.

Table 1
Fractional atomic positions and thermal parameters used in the *CRYSTAL*95 calculations (all values × 10^4).

	x	y	z	$U_{11} = U_{22}$	U_{33}	U_{12}	$U_{13} = U_{23}$
C	0	5000	3283	167	86	−1	0
O	0	5000	5963	220	79	24	0
N	1447	x + 5000	1784	317	113	−163	3
H_1	2552	x + 5000	2845	484	183	−325	−15
H_2	1428	x + 5000	−339	486	113	−303	21

O κ parameters were introduced to allow for expansion/contraction of the valence shell (Coppens *et al.*, 1979). The inner *K*-shells were assigned fixed populations of 2 electrons for C, N and O, whereas the number of valence-shell electrons was unconstrained. The radial scattering factor for H as well as the multipole functions for C, N, O and H were obtained from single Slater-type density functions with standard values for the radial exponents [α = 6.43, 7.37, 8.51 and 4.69 Å^{-1} (Hehre *et al.*, 1969)].


(*a*)

(*b*)

Figure 1
Deformation densities calculated using the *CRYSTAL*95 (Dovesi *et al.*, 1996) program. (*a*) Hartree–Fock and (*b*) DFT. Positive contours (electron excess) are drawn as solid lines, zero contours are dash-dotted and negative contours (charge deficiency) are dotted. Contour intervals are at 0.01 e a.u.$^{-3}$.

3. Calculations

3.1. Quantum chemical calculations

The EDD of the urea crystal, space group $P\bar{4}2_1m$, has been calculated in the geometry that was experimentally obtained by Zavodnik *et al.* (1999) at 148 K, *a* = 5.589 and *c* = 4.6947 Å. The positions of the atoms that were used are presented in Table 1.

Hartree–Fock and density functional calculations were performed in this geometry, both in the $6-21G^{**}$ basis set. Whereas this basis is small for molecular calculations, it is quite large for solid-state ones owing to the basis set superposition. There are several choices for the exchange and correlation potential in the DFT calculation. We have chosen the exchange potential according to Becke (1988; BECKE) and the correlation potential in the generalized gradient approximation according to Perdew & Wang (1986, 1989, 1992; PWGGA). A DFT calculation with the correlation part modeled according to Vosko *et al.* (1980) in the local density approximation resulted in a EDD that was extremely close to *BECKE–PWGGA* results. We therefore will continue presenting only the *BECKE–PWGGA* calculations.

The deformation densities, which are defined as the crystal densities from which the superposition of non-interacting atoms has been subtracted, derived from HF and DFT calculations, are presented in Figs. 1(*a*) and (*b*). The *CRYSTAL*95 program uses the atomic program developed by Roos *et al.* (1968) for the calculation of the atomic densities, so we have to keep in mind that from both total EDDs, obtained with HF and DFT, respectively, the same superposition of atomic densities was subtracted.

The deformation densities resulting from the HF and DFT calculations are remarkably similar. The accumulation of charge in the bonding areas differs by no more than one contour. The only striking difference is the difference in density of the π-electrons in the carbon region. The resulting interaction density (only for the DFT calculation) is shown in Fig. 2.

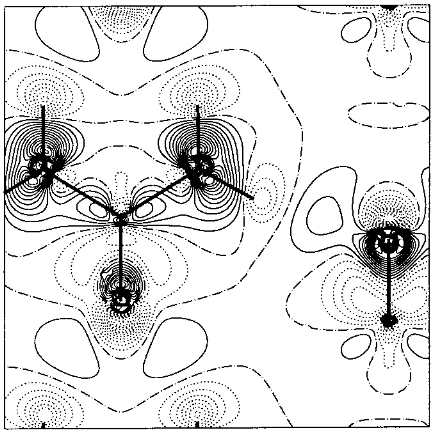


Figure 2
The interaction density calculated using DFT. Contour intervals are at 0.002 e a.u.$^{-3}$.

Structure factors can be calculated from both EDDs. In crystallography, the agreement between two datasets is usually expressed in the form of an R factor, $R = \Sigma||F_{\rm obs}| - |F_{\rm calc}||/\Sigma|F_{\rm obs}|$ or a weighted R factor, $wR = \left(\Sigma w(|F_{\rm obs}| - |F_{\rm calc}|)^2/\Sigma w|F^2_{\rm obs}|\right)^{1/2}$. It is the latter expression that is minimized in a refinement. The weight w is in most cases taken as $1/\sigma^2(F_{\rm obs})$, where $\sigma(F_{\rm obs})$ is the standard deviation of the observed structure factor. The dynamic structure factors that take the place of $F_{\rm obs}$ are obtained from the theoretical, static structure factors as described in the previous section. The thermal parameters that were used in this procedure are also shown in Table 1. The weights were derived from the standard deviations that were obtained in the experiment of Zavodnik *et al.* (1999), who measured 412 reflections at 148 K. The resulting R factor calculated from the static data is 0.0058. From the comparison of the dynamic data we obtain R = 0.0070 and wR = 0.0111.

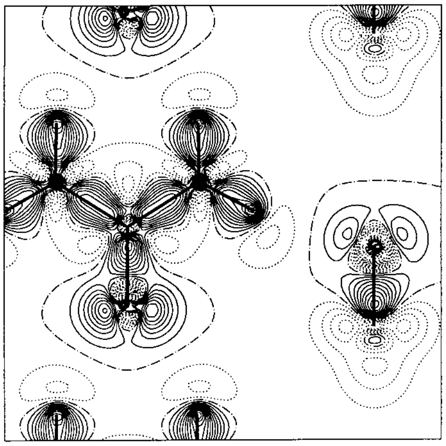


Figure 3
The deformation density after refinement of the static data. Contour intervals are at 0.01 e a.u.$^{-3}$

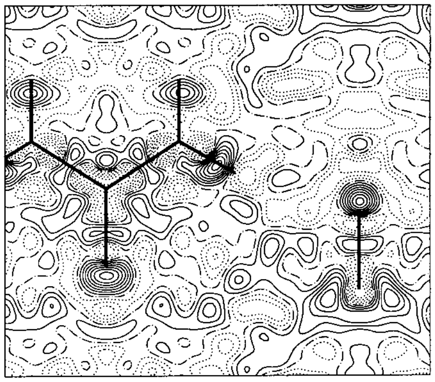


Figure 4
The difference between the residual density, where the phase of the multipole model was assigned to the observed structure factors, and the residual density, where the phase of the DFT calculation was assigned to the observed structure factors. Contour intervals are at 0.002 e a.u.$^{-3}$.

3.2. Multipole refinement

To check whether the multipole model is capable of modeling the EDD we selected the DFT dataset to use in the refinement procedure. A refinement of the static data was performed with all the thermal parameters set to zero. The scale factor was fixed at 1. Unit weights were assigned to the structure factors.

All multipole parameters up to hexadecapole were included for the C, N and O atoms, and up to quadrupole for the H atoms. κ parameters of the C, N and O atoms were also refined. This resulted in a total number of 46 parameters. The agreement in terms of R and wR is $R = 0.60\%$ and $wR = 0.64\%$. After this refinement the values of the multipoles that were smaller than their standard deviation were set to zero. This reduced the number of parameters to 34, with $R = 0.61\%$ and $wR = 0.64\%$. The resulting deformation density is shown in Fig. 3. Small differences are visible between Fig. 3 and Fig. 1(*b*), particularly in the core regions, but the main characteristic features in the (hydrogen) bonding regions persist.

The density that could not be modeled can be shown in a residual map, a Fourier summation of the difference between observed and calculated structure factors. However, in the case of a non-centrosymmetric structure, the presented

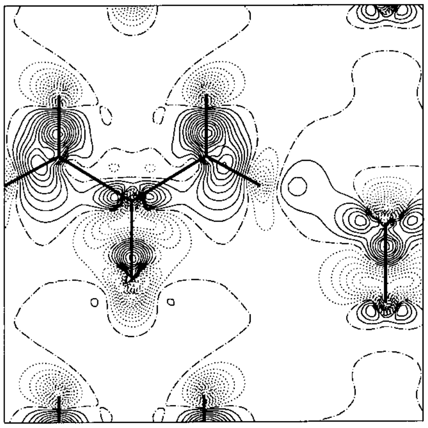


Figure 5
The interaction density after refinement of the static data. Contour intervals are at 0.002 e a.u.$^{-3}$.

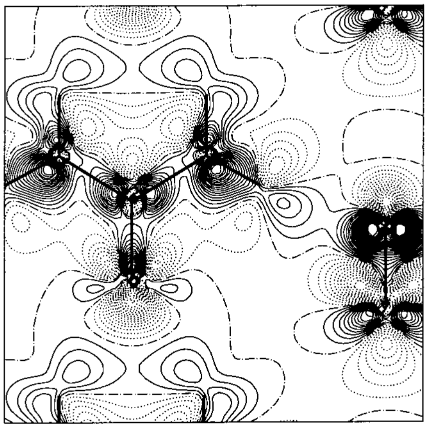
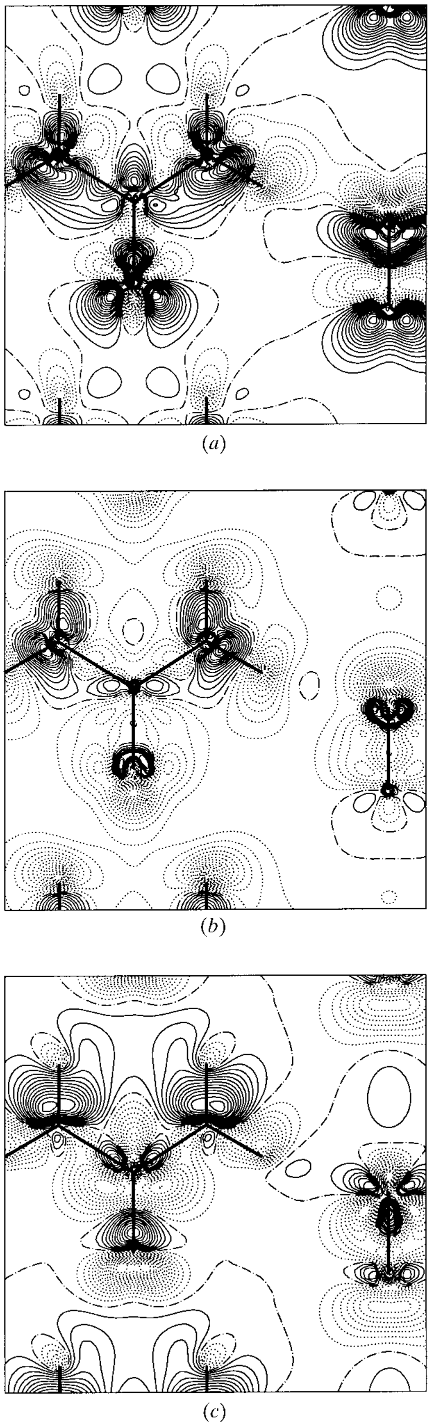


Figure 6
The interaction density after refinement of the dynamic data where noise had been added. Contour intervals are at 0.002 e a.u.$^{-3}$.

density does not reflect the real residual density. In calculating the difference between observed and calculated structure factors, the phase of the *calculated* structure factors is assigned to the *observed* structure factors, because this phase is not known. In the case of our model calculation the correct phase is known. We have calculated the residual densities both ways (with correct and multipole phases) and subtracted them. The result is shown in Fig. 4. It reflects the electron density that was not accounted for owing to loss of phase information. We can also account for this effect in the calculation of the R factors leading to $R = 0.81\%$ and $wR = 0.82\%$. Fig. 4 shows clearly that the electron density which could not be modeled owing to the loss of phase information is of the same order of magnitude as the interaction density. The error is clearly due to the fact that the structure factors based on the modeled EDD have a slightly deviating phase.

Table 2
The refinement statistics of four different random noise distributions (weighted R factor, R factor and goodness-of-fit).

Dataset	*a*	*b*	*c*	*d*
wR	1.47	1.40	1.38	1.44
R	1.11	1.00	1.12	1.09
GOF	1.28	1.22	1.21	1.25


(*a*)

(*b*)

(*c*)

Figure 7
The interaction densities calculated from various dynamic data sets (different random noise was added). Contour intervals are at 0.002 e a.u.$^{-3}$.

To obtain the interaction density we have to subtract the sum of isolated molecules from the experimentaly determined electron density distribution. If systematic errors are introduced by the multipole refinement procedure we can try to eliminate these by applying the same multipole refinement to the diffraction data of the sum of isolated molecules, where we can expect the same systematic errors. The result of such a calculation on static data is shown in Fig. 5. It is clear that this time all the characteristic features of the interaction density are preserved in this picture. When the same procedure is applied to the dynamic data using the experimental weights and only refining multipole parameters we also find a preservation of the characteristic features.

3.3. Addition of noise

However, in an experimental situation we always have noise on the data. It is the intention that the multipole model will filter out this noise. We have simulated noise by multiplying each standard deviation with a random number selected from a Gaussian distribution with unit variance and adding this to the corresponding structure factor. Following the same procedure as outlined above (no noise was added to the data calculated from the sum of isolated molecules), we obtain the result shown in Fig. 6. We see that many features that were present in the original interaction density are heavily distorted or have disappeared completely. By no means a reliable interaction density can be identified.

To confirm the effect of noise on the data we have calculated three more noisy data sets, each using a different sequence of random numbers, and carried out refinements on each of these data sets. The results of these refinements is shown in Fig. 7. The differences in these resulting electron density distributions show that it is nearly impossible to extract the effect of intermolecular interaction from diffraction data. The resulting refinement statistics are presented in Table 2. The values of the goodness-of-fit are all slightly higher than one, which is in agreement with our expectations as the model cannot describe the density completely (the residual density of the refinements of the noiseless data is not zero).

4. Discussion and conclusion

In this paper we have demonstrated that it is not possible to extract the effect of intermolecular interaction in urea from diffraction data with the current multipolar refinement techniques. We have used simulated diffraction data, calculated from a known EDD, to show this. Although in this case the scale factor, positional and thermal parameters, cell dimensions and error distribution of the noise on the data were known exactly, it was not possible to retrieve the interaction density. This means that in a real experiment, where these parameters are known only approximately, the results will be worse.

We are aware of the fact that our model study only refers to the particular case of urea and that an experiment in which more structure factors are measured more accurately would give better results. However, we believe that the experiment we have taken as a reference for our model study has been carried out with great care and we expect our problems to occur in similar high quality experiments on non-centrosymmetric crystals.

References

Becke, A. D. (1988). *Phys. Rev. A*, **38**, 3098–3103.

Bruning, H. & Feil, D. (1992). *Acta Cryst.* A**48**, 865–872.

Clementi, C. & Roetti, C. (1974). *At. Data Nucl. Data Tabl.* **14**, 177.

Coppens, P., Guru Row, T. N., Leung, P., Stevens, E. D., Becker, P. J. & Yang, Y. W. (1979). *Acta Cryst.* A**35**, 63–72.

Craven, B. M., Weber, H. P. & He, X. (1987). Technical Report TR-87-2. Department of Crystallography, University of Pittsburgh, PA 15620, USA.

Cromer, D. T. & Waber, J. T. (1974). *International Tables for X-ray Crystallography*, Vol. IV, pp. 71–147. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)

Dovesi, R., Saunders, V. R., Roetti, C., Causà, M., Harrison, N. M., Orlando, R. & Aprà, E. (1996). *CRYSTAL95 User's Manual.* University of Torino, Italy.

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

Hermansson, K. (1984). PhD Thesis. University of Uppsala, Upssala, Sweden.

Hehre, W. J., Stewart, R. F. & Pople, J. A. (1969). *J. Chem. Phys.* **51**, 2657–2664.

Hirshfeld, F. L. (1971). *Acta Cryst.* B**27**, 769–781.

Hirshfeld, F. L. (1977). *Isr. J. Chem.* **16**, 226–229.

Koritsansky, T., Hansen, N., Howard, S., Su, Z. & Mallinson, P. (1994). *The XD Program Package*. First release May 1994. Berlin, Glasgow.

Perdew, J. P. & Wang, Y. (1986). *Phys. Rev. B*, **33**, 8800–8802.

Perdew, J. P. & Wang, Y. (1989). *Phys. Rev. B*, **40**, 3399.

Perdew, J. P. & Wang, Y. (1992). *Phys. Rev. B*, **45**, 13244–13249.

Roos, B., Salez, C., Veillard, A. & Clementi, E. (1968). Technical Report RJ518 IBM Res.

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

Spackman, M. A., Byrom, P. G., Alfredsson, M. & Hermansson, K. (1999). *Acta Cryst.* A**55**, 30–47.

Stewart, R. F. & Spackman, M. A. (1983). *Valray User's Manual.* Department of Chemistry, Carnegie-Mellon University, Pittsburgh, PA, USA.

Swaminathan, S., Craven, B. M. & McMullan, R. K. (1984). *Acta Cryst.* B**40**, 300–306.

Vosko, S. H., Wilk, L. & Nusair, M. (1980). *Can. J. Phys.* **58**, 1200–1211.

Zavodnik, V. E., Stash, A. I., Tsirelson, V. G., de Vries, R. Y. & Feil, D. (1999). *Acta Cryst.* B**55**, 45–54.