Response to the paper *A comparison between experimental and theoretical aspherical-atom scattering factors for charge-density refinement of large molecules,* by Pichon-Pesme, Jelsch, Guillot & Lecomte (2004)

Anatoliy Volkov, Tibor Koritsanszky, Xue Li and Philip Coppens

letters to the editor

Response to the paper *A comparison between experimental and theoretical aspherical-atom scattering factors for charge-density refinement of large molecules,* by Pichon-Pesme, Jelsch, Guillot & Lecomte (2004)

Anatoliy Volkov,[a]* Tibor Koritsanszky,[b]* Xue Li[a] and Philip Coppens[a]*

[a]Department of Chemistry, State University of New York at Buffalo, Buffalo, NY 14260-3000, USA, and [b]Department of Chemistry, Middle Tennessee State University, MTSU Box 0395, 1301 East Main Street, Murfreesboro, TN 37132, USA. Correspondence e-mail: volkov@chem.buffalo.edu, tkoritsa@mtsu.edu, coppens@buffalo.edu

A paper was recently published in *Acta Crystallographica* Section A by Pichon-Pesme *et al.* (2004), called PJGL in the following, that deals with the comparison of the *experimental* databank of aspherical pseudoatom parameters developed by those authors with the relatively new *theoretical* approach reported recently by us (Koritsanszky *et al.*, 2002).

As conclusions, PJGL list that:

(i) 'the electron-density maps derived from the two methods [theoretical and experimental] agree qualitatively only';

(ii) '[the difference in population parameters] leads to different electrostatic properties';

(iii) 'the charges of the experimental database are in accordance with chemical intuition';

(iv) 'the experimental database has been successfully tested on protein high-resolution X-ray data'.

For a reader not very familiar with the field, these conclusions clearly imply that the theoretical approach is rather inferior to the experimental one. The PJGL paper states that the electrostatic potential derived from theoretical pseudoatoms is similar to that from a 'semi-empirical molecular-orbital calculation'. However, a quantitative comparison with more advanced first-principles calculations is more appropriate.

We have recently published a series of papers (Volkov & Coppens, 2004; Volkov, Li *et al.*, 2004; Volkov, Koritsanszky & Coppens, 2004) in which the principles of construction of the theoretical databank are described and comparisons with first-principles calculations, such as the density functional theory (DFT) (Hohenberg & Kohn, 1964; te Velde *et al.*, 2001), are made, using a large number of electron-density (ED) related properties. These include (i) the total ED, $\rho(\mathbf{r})$; (ii) the deformation density; (iii) the electrostatic potential; (iv) a number of topological properties of the ED (Bader, 1990) [location of critical points, properties of $\rho(\mathbf{r})$ at the critical points – density, curvatures, Laplacian, ellipticity, topological net atomic charges and higher atomic moments up to the hexadecapolar level]; (v) total molecular moments; and, more importantly, (vi) the intermolecular electrostatic interaction energies, including electrostatic binding energies of molecules in crystals.

The level of accuracy in the prediction of ED properties by the theoretical databank is found to be comparable with that from various first-principles calculations with different Hamiltonians and basis-set expansions of the wavefunction (Volkov, Li *et al.*, 2004). As an example, the bonding features of the deformation density maps are predicted within 0.02–0.1 e Å^{-3}, while the magnitude of $\rho(\mathbf{r})$ at

critical points is reproduced well within 0.1 e Å^{-3} when compared to a rather extensive B3LYP/6-311++G(3df,3pd) calculation. Discrepancies of similar magnitude are observed, for example, between B3LYP/6-311++G(3df,3pd) and BLYP/TZP calculations. The agreement of the databank-derived topological properties of the ED is noticeably improved when comparison is made with those from DFT calculations in which the Slater-type radial functions (triple-zeta basis, TZP) are used in the basis-set expansion. In fact, significant discrepancies in second derivatives of $\rho(\mathbf{r})$ at the bond critical points between theory and experiment, which in the past have often been attributed to the limited flexibility of the density functions in the pseudoatom model (see, for example, Swaminathan *et al.*, 1984, and Volkov & Coppens, 2001), are found to be at least in part due to the different behaviors of the Gaussian- and Slater-type radial functions in the vicinity of critical points (Volkov, Li *et al.*, 2004).

In our first studies (Volkov & Coppens, 2004; Volkov, Li *et al.*, 2004), the agreement in electrostatic interaction energies in molecular dimers between DFT calculations, obtained using the Morokuma–Ziegler energy decomposition scheme (Morokuma, 1971; Ziegler & Rauk, 1977), and the theoretical databank was not very impressive, the root-mean-squared (r.m.s.) deviation from the DFT/BLYP/TZP electrostatic interaction energies being about 25 kJ mol^{-1}. However, these discrepancies were due to the limitation of the multipolar expansion of the electrostatic interaction energy (Buckingham, 1967) and *not* the problem of the theoretical databank. Subsequent application of the newly developed exact potential and multipole method (EP/MM) (Volkov, Koritsanszky & Coppens, 2004) to the theoretical databank parameters significantly reduces the discrepancies to at most 20 kJ mol^{-1} (typically below 10 kJ mol^{-1}) and the r.m.s. discrepancy to only 8 kJ mol^{-1} !

The comparison of individual pseudoatom parameters performed by PJGL is not relevant as the individual parameters do not have a well defined physical meaning. The PJGL comment that 'the charges of the experimental database are in accordance with chemical intuition' is a rather ill defined statement. A valid test can only be provided by the physical properties calculated from the pseudoatom expansion of the ED. While this point seems to be accepted by PJGL (see §3 of Pichon-Pesme *et al.*, 2004), much of their paper is still dedicated to the comparison of individual pseudoatom parameters. However, the performance of the theoretical databank has been verified by the thorough comparison with extensive density functional calculations. On the other hand, while the original PJGL databank dates back as far as 1995 and (as stated) is now complete, to

our knowledge, no such studies using the experimental databank have been published yet, with the exception of some deformation density maps (Pichon-Pesme *et al.*, 1995; Jelsch *et al.*, 2000) and several two-dimensional maps of the electrostatic potential (Muzet *et al.*, 2003; Pichon-Pesme *et al.*, 2004). In practice, the latter can be reasonably well predicted even by the point-charge model (Williams, 1988; Breneman & Wilberg, 1990).

Given the good fit with theory of the results obtained with the theoretical databank and the conclusion (ii) that different electrostatic properties are expected from different population parameters, similar tests of the experimental databank appear highly desirable.

We also point out the excellent internal consistency of the theoretical databank, as the error in prediction of the total molecular electroneutrality of several amino acid and peptide molecules, including Leu-enkephalin, was found to be well under 1% (usually <0.5%) (Volkov, Li *et al.*, 2004).

Finally, the extension of the theoretical databank to include additional atom types is straightforward and can be done in a matter of days, compared to the more time-consuming experimental procedure, which is also subject to thermal smearing effects and possible phase ambiguities.

In conclusion, we would like to emphasize that it is not the intent of this letter to criticize the experimental databank, which fulfils a useful function and has indeed been found to be a good starting point for aspherical refinements of macromolecules (Jelsch *et al.*, 1998, 2000; Fernández-Serra *et al.*, 2000; Muzet *et al.*, 2003). However, before any conclusions can be drawn on whether or not the experimental databank is superior to the theoretical one, it must be tested in the same manner that the theoretical databank has been scrutinized.

References

Bader, R. F. W. (1990). *Atoms in Molecules. A Quantum Theory*. New York: Oxford University Press.

Breneman, C. M. & Wilberg, K. B. (1990). *J. Comput. Chem.* **11**, 361–373.

Buckingham, A. D. (1967). *Adv. Chem. Phys.* **12**, 107–142.

Fernández-Serra, M. V., Junquera, J., Jelsch, C., Lecomte, C. & Artacho, E. (2000). *Solid State Commun.* **116**, 395–400.

Hohenberg P. & Kohn, W. (1964). *Phys. Rev.* **136**, B864–B871.

Jelsch, C., Pichon-Pesme, V., Lecomte, C. & Aubry, A. (1998). *Acta Cryst.* D**54**, 1306–1318.

Jelsch, C., Teeter, M. M., Lamzin, V., Pichon-Pesme, V., Blessing, R. H. & Lecomte, C. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 3171–3176.

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

Morokuma, K. (1971). *J. Chem. Phys.* **55**, 1236–1244.

Muzet, N., Guillot, B., Jelsch, C., Howard, E. & Lecomte, C. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 8742–8747.

Pichon-Pesme, V., Jelsch, C., Guillot, B. & Lecomte, C. (2004). *Acta Cryst.* A**60**, 204–208.

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

Swaminathan, S., Craven, B. M., Spackman, M. A. & Stewart, R. F. (1984). *Acta Cryst.* B**40**, 398–404.

Velde, G. te, Bickelhaupt, F. M., van Gisbergen, S. J. A., Fonseca Guerra, C., Baerends, E. J., Snijders, J. G. & Ziegler, T. J. (2001). *J. Comput. Chem.* **22**, 931–967.

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

Volkov, A. & Coppens, P. (2004). *J. Comput. Chem.* **25**, 921–934.

Volkov, A., Koritsanszky, T. & Coppens, P. (2004). *Chem. Phys. Lett.* **391**, 170–175.

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

Williams, D. E. (1988) *J. Comput. Chem.* **9**, 745–763.

Ziegler, T. & Rauk, A. (1977). *Theor. Chim. Acta*, **46**, 1–10.