

Comparative refinement of correct and incorrect structural models of tetrabutylammonium tetrabutylborate – pitfalls arising from poor-quality data

Vladimir Stilinović and Branko Kaitner*

Department of Chemistry, Faculty of Science, University of Zagreb, Horvatovac 102a, 10000 Zagreb, Croatia. Correspondence e-mail: kaitner@chem.pmf.hr

This paper demonstrates how numerical parameters usually used to assess the quality of a crystal structure solution (R , wR and S) may be misleading when studying a model refined against poor-quality data. Weakly diffracting crystals of tetrabutylammonium tetrabutylborate, a low-density organic salt comprising isoelectronic cations and anions, were measured using Cu and Mo $K\alpha$ radiation. Along with the correct structural model, six erroneous structural models were constructed and refined against the same data. For both data sets it was found that models based on an incorrect unit-cell choice give lower values of R and wR than the correct one, thus apparently being in better agreement with measured data. Closer inspection of the measured data shows that this is in fact not the case.

© 2010 International Union of Crystallography
Printed in Singapore – all rights reserved

1. Introduction

Over the past five decades, X-ray diffraction on crystalline materials has become the basic method for crystallographic research. The development of apparatus for data collection on the one hand and software for crystal structure determination and refinement on the other has resulted in a great increase in the number of crystal structures determined every year. Unfortunately, this increase in number is accompanied by a no less notable increase in crystal structures which have been incorrectly determined. Among the most common errors are incorrect space-group and unit-cell assignment (Marsh, 1995, 2002; Marsh *et al.*, 2002; Marsh & Herbstein, 1988; Herbstein & Marsh, 1998; Clemente & Marzotto, 2004), but there are also cases of incorrectly assigned atom types, missing or incorrectly placed H atoms and other errors (Spek, 2003, 2009). When atom types are wrongly assigned, it is usually due to confusion of (approximately) isoelectronic atoms such as mistaking oxygen for fluorine, or silicon for chlorine (von Schnering & Vu, 1983) due to the similarity of their scattering powers.

To study potential problems with small-molecule structure determination we have designed a particularly difficult crystal. The substance chosen was tetrabutylammonium tetrabutylborate, (I), an organic salt with a total of 34 non-H atoms per formula unit. On the one hand, the thermal motion of the butyl groups will cause an irresolvable disorder which will reduce the quality of the model. Furthermore, the absence of hydrogen bonding and other directional interactions between the moieties in the crystal means that the crystal packing is expected to be quite loose, the crystal density low and the mosaicity high, which will decrease the diffraction intensity,

and thus the quality of the X-ray data collected. On the other hand, since the cation and the anion are both of the same geometry and isoelectronic, they should be quite difficult to distinguish using X-ray diffraction.

2. Experimental

2.1. Preparation of the compound

A hexane solution of butyllithium (0.20 mol in 30 ml) was added dropwise into a diethylether solution of boron trifluoride (0.04 mol in 20 ml) for 15 min. After all the butyllithium was added, the solution was heated for 1 h, after which the solvent was evaporated and the remaining slurry dissolved in hot water. Hot aqueous solution of tetrabutylammonium iodide (0.04 mol in 25 ml) was added, leading to the immediate precipitation of (I), which was crystallized twice from diisopropylether. Crystals suitable for diffraction experiments were grown by slow evaporation of an ethanol solution at 268 K over approximately 1 month.

2.2. Crystal measurement and structure refinement

The diffraction data were collected at 295 (2) K for both measurements.¹ One data set was collected using an Oxford Diffraction Xcalibur Kappa CCD X-ray diffractometer with graphite-monochromated Mo $K\alpha$ (0.71073 Å) radiation. The other data set was collected using an Oxford Diffraction Xcalibur Nova R, using graphite-monochromated microfocus tube Cu $K\alpha$ (1.54179 Å) radiation. Data were reduced using

¹ Supplementary data for this paper are available from the IUCr electronic archives (Reference: ZM5071). Services for accessing these data are described at the back of the journal.

the *CrysAlis RED* program suite (Oxford Diffraction, 2003), and the structures were solved and refined using the *SHELXS* and *SHELXL* programs (Sheldrick, 1997, 2008). The H atoms were placed in calculated positions and treated as riding on their parent atoms. The packing diagram was prepared using *ORTEP-3* (Farrugia, 1997) and *POVRay* (Persistence of Vision Pty, 2004).

3. Structural models

The measurement using Mo radiation indicated that the unit cell is tetragonal [$P4_2/n$, $a = 13.393(5)$, $c = 5.300(5)$ Å] with one formula unit per unit cell. Although it is possible to solve the structure based on this unit-cell choice and even to refine it to quite a reasonable degree, such a structural model (which for convenience we shall refer to as structural model *B*) cannot be correct. The symmetry of the unit cell requires the two ions constituting the formula unit to be equivalent. This is only possible if the structure is disordered, *i.e.* the cations and the anions are positioned randomly throughout the structure, which is both contrary to our previous knowledge about structures of ionic salts and to chemical and physical common sense. Furthermore, the structures of similar compounds (MPh_4BPh_4 , $M = P, As, Sb$; Lloyd & Brock, 1997) are all ordered, tetragonal (space group $I\bar{4}$), with four formula units per unit cell and ions alternating along the a and c directions. Therefore it is necessary to model the structure of (I) *via* a

reasonable unit cell. It can be obtained from the original one by the use of the transformation matrix $T = (1, 1, 0; \bar{1}, 1, 0; 0, 0, 2)$ which leads to a tetragonal unit cell [$I4_1/a$, $a = 18.958(5)$, $c = 10.624(5)$ Å], with the volume and Z increased by a factor of 4 (Fig. 1). The structure solution and refinement according to this model led to a reasonable model with alternating symmetry-independent cations and anions. Nitrogen and boron atoms could be correctly assigned according to bond lengths since the distance between the central atom and the first butyl carbon was found to be 1.512(2) Å [Cambridge Structural Database (CSD) (Allen, 2002) average for N–C in tetrabutylammonium is 1.522 Å] in one ion and 1.654(2) Å (CSD average for B–C in all tetraalkylborates is 1.648 Å, there being only one entry for the tetrabutylborate anion) in the other. Thus the correct structural model (model *A*) was obtained.

Another crystal from the same batch was re-measured using Cu radiation, which confirmed the cell choice, the series of weak reflections which were lost in the background radiation in the measurements with Mo $K\alpha$ radiation now being obvious (Fig. 2). The solution and the refinement of the structure based on these data gave the structural model *A*, with only minor differences in bond lengths [N1–C1 distance of 1.547(5) Å and B1–C5 distance of 1.635(5) Å], but with a lower R of only 0.0643, compared to 0.0807 obtained from Mo radiation

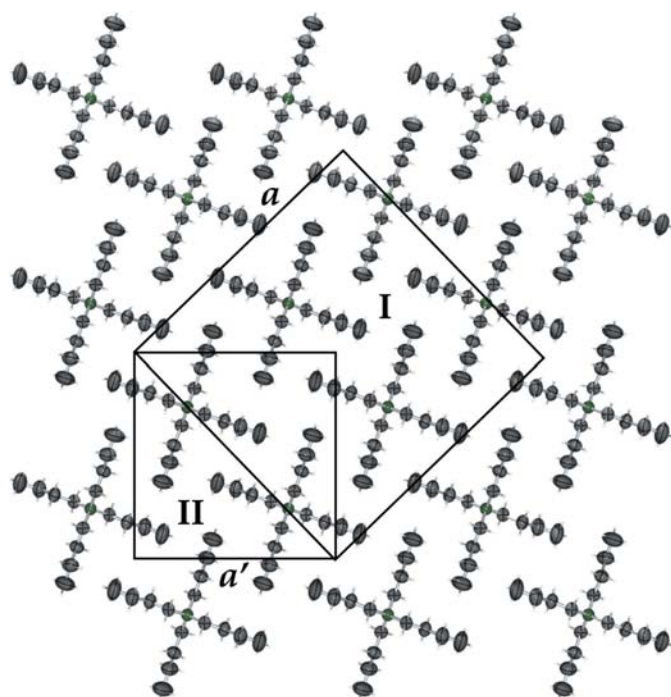


Figure 1

The view of the structure of tetrabutylammonium tetrabutylborate depicting two unit-cell choices – I, the actual cell (the basis for models *A*, *C–F*) and II, the subcell (the basis for models *B*, *G* and *H*) – and the respective unit-cell parameters a and a' . Displacement ellipsoids are shown with 30% probability and H atoms are depicted as spheres of arbitrary radii.

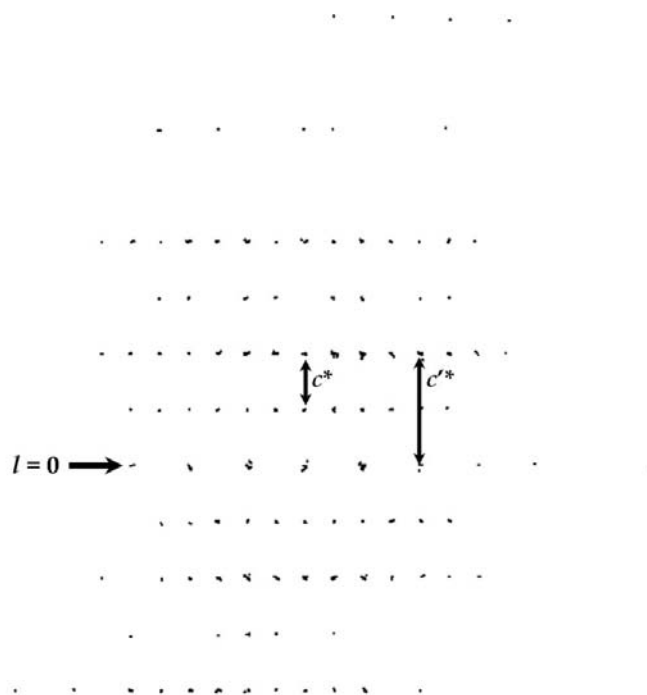


Figure 2

Measured reflections (Cu radiation) plotted in the reciprocal space viewed along a^* (in the correct cell choice), showing alternate layers of strong and weak reflections. If weak reflections are omitted, a smaller unit cell is obtained (cell choice II, corresponding reciprocal-lattice parameter c'^* shown). If all reflections are used, the correct unit cell is obtained (cell choice I, corresponding reciprocal-lattice parameter c^* shown). The central layer ($hk0$) also shows systematic absences (k odd) indicating the presence of an a -glide plane.

Table 1

Statistical descriptors of data and structural model quality for the correct (*A*) and seven incorrect (*B–H*) structural models with comparison between the data obtained by diffraction of Mo and Cu radiation.

	R_{obs}		R_{all}		wR_{obs}		wR_{all}		S	
	Cu	Mo	Cu	Mo	Cu	Mo	Cu	Mo	Cu	Mo
<i>A</i>	0.0643	0.0807	0.1793	0.3277	0.2165	0.1944	0.2713	0.2554	0.723	0.744
<i>B</i>	0.0604	0.0681	0.1212	0.2383	0.1692	0.1800	0.2019	0.2439	0.762	0.915
<i>C</i>	0.0850	0.0893	0.2192	0.3122	0.2824	0.2177	0.3795	0.2783	0.877	0.772
<i>D</i>	0.1100	0.1102	0.2270	0.2989	0.3353	0.3358	0.4094	0.4021	0.939	0.863
<i>E</i>	0.0762	0.0912	0.1883	0.3105	0.2341	0.2215	0.2952	0.2881	0.760	0.751
<i>F</i>	0.0969	0.1047	0.2239	0.3321	0.2889	0.2631	0.3784	0.342	0.849	0.681
<i>G</i>	0.0601	0.0647	0.1205	0.2106	0.1714	0.1722	0.2036	0.2117	0.768	0.788
<i>H</i>	0.0590	0.0575	0.1286	0.2080	0.1641	0.1417	0.2091	0.1754	0.750	0.667

(Table 1). However, if a subcell analogous to that obtained by Mo radiation is chosen (obtained by multiplying the cell by \mathbf{T}^{-1}) and the data reduced with respect to such a unit-cell choice, structure solution yields structural model *B*, which can also be quite reasonably refined. This fact prompted us to perform further refinements of erroneous structural models, namely: model *C*, disordered model (random distribution of ions in the actual unit cell); model *D*, positions of boron and nitrogen atoms interchanged; model *E*, nitrogen and boron atoms replaced by carbon; model *F*, non-centrosymmetric structure (space group $\bar{1}4$); model *G*, central atom replaced by carbon using the subcell as in model *B*; and model *H*, non-centrosymmetric structure (space group $P\bar{4}$), using the subcell as in models *B* and *H* (Fig. 3).

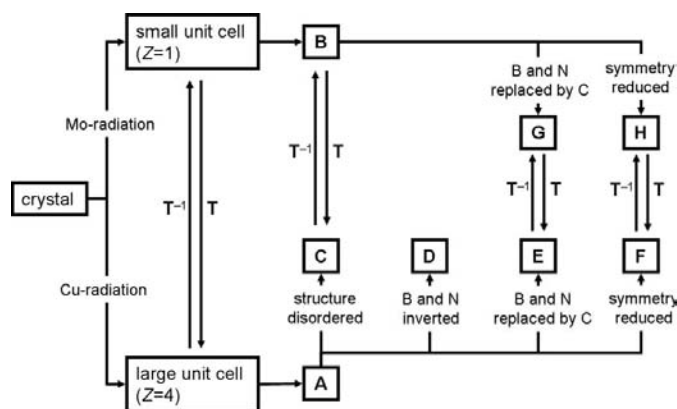
3.1. Comparison of the refined structural models

All these structural models were refined without any problems and they are all quite similar to one another. According to all models the structure comprises cross-like molecules of S_4 symmetry (or approximate S_4 symmetry in the cases of models *F* and *H*) with thermal ellipsoids increasing

towards the terminal methyl C atoms of the butyl chains arranged so that the closest neighbours are packed in columns along the *c* axes. Every such column is surrounded by four others so that the *z* coordinate of any molecule is (approximately for model *F*, otherwise exactly) equal to the average of the *z* coordinates of its nearest neighbours in neighbouring columns. All models correspond to low-density structures (*ca* 0.85 g cm^{-3}). The low density of the crystal is, to an extent, also responsible for weak diffraction of the sample (unusually low ratio of observed to measured reflections).

Statistical parameters significant for data and model quality (*R* and weighted *R* for observed and all data and the goodness of fit, *S*) are given in Table 1. Refinement of three models based on the actual symmetry and the correct unit cell, namely the disordered model *C*, the model with interchanged boron and nitrogen (*D*) and the model with all central atoms replaced by carbon (*E*) as well as the non-centrosymmetric model *F*, gives less favourable values of the above parameters than the correct model. However, it is not quite possible to determine which would be the ‘worst’ model. While model *D* gives the largest values of *R* and *wR* for both Cu and Mo radiation, *S* would indicate the opposite, being closer to unity for model *D* than for *C*. It is also important to note that model *C* failed to converge completely when refined against the Cu data with the maximum shift/s. u. ratio of *ca* 7.5 even after 400 refinement cycles.

Non-centrosymmetric model *F*, although having a lower *R* than model *D*, was shown to have a number of other, quite serious, problems. Refinement of the model *F* against both Cu and Mo radiation data shows it to be quite inferior to the correct model. Both *R* and *wR* are quite a bit higher, only *S* is closer to unity than for the model *A*. There are, however, several other problems with this model. The U_{eq} values for a number of atoms strongly deviate from their neighbours and displacement ellipsoids are somewhat suspicious. Also, unlike most of the other models, *F* fails to properly converge with the maximal shift/s. u. ratio remaining above 3 for both data sets (3.148 for Cu and 5.519 for Mo), even after more than 2000 least-squares refinement cycles. This is due to the impossibility of methyl H atoms (modelled as idealized methyl groups with torsion angles determined from electron density) settling in one set of positions. In addition, the *ADDSYM* application of *PLATON* (Spek, 1998) strongly suggests that the structure is

**Figure 3**

Schematic representation of the derivation and interrelations of the correct structural model *A* and the incorrect models *B–H* with noted possible transformations from small unit-cell to respective large unit-cell models and *vice versa* by multiplication with transformation matrix \mathbf{T} or its inverse \mathbf{T}^{-1} .

centrosymmetric. The wrong space group is also obvious from a closer inspection of systematic absences, which quite clearly indicate the presence of *a*-glide planes perpendicular to *c* (Fig. 2), which are not present in the space group $\bar{I}4$.

A more dramatic effect is noticeable using structural models *B*, *G* and *H*, which are based on the incorrectly chosen unit cell. For all three models *R* and *wR* are significantly lower than for the correct model *A*, and *S* for models *B* and *G* is closer to unity than for model *A* (for respective data sets). Unlike model *F*, full convergence of the models is quite easily achieved, there are no anomalous atomic displacement parameters and the displacement ellipsoids are quite reasonable. The bond lengths between the central atom and the first butyl carbon in *B* and *G* are intermediate to N–C and B–C bond lengths in *A* [model *B*, Cu 1.5790 (17) Å, Mo 1.5785 (18) Å; model *G*, Cu 1.5802 (16) Å, Mo 1.5804 (18) Å]. The relative positions of the molecules in *B* are virtually identical to those in model *C*, in *G* to *E* and in *H* to *F*.

Although the fact that incorrect models can appear to be much ‘better’ than the correct model is surprising, the apparent higher quality of the models based on the smaller unit cell is in fact rather easy to explain. One needs to keep in mind the poor quality of the data, *i.e.* quite low ratio of observed [$I > 2\sigma(I)$] to measured reflections. Using the smaller cell, the number of reflections used in data reduction and subsequent crystal structure solution and refinement is one quarter of the overall number of reflections used when data reduction is based on the actual cell. The remaining reflections are those which are systematically absent due to the cell centring (one half of the overall number of reflections) and the weak reflections which correspond to reflections of type *hkl*, where *l* is odd if the correct unit cell is used (the remaining quarter of the overall number of reflections). Since these reflections were quite weak (for Mo radiation almost entirely unobserved), they are measured with substantial error. Removing them from the data set also removes the error of their measurement and increases the ratio of observed to measured reflections, thus making the data appear to be of higher quality. The most obvious indication of this effect is the large difference between the *R* and *wR* calculated for observed and for all measured reflections (the difference being larger for models based on the correct unit cell than for models based on the smaller cell; Table 1).

According to the parameters given in Table 1, model *H* appears to be the best structural model and model *G* the second best. When refined against Mo data both *R* and *wR* are lower for *H* than for any other model. When refined against Cu data *R* for all and *wR* for observed reflections are insignificantly higher than for *B*. Only *S* deviates from unity more for *H* than for most other models but for Cu data it is closer to unity than for the correct model *A*. Note that, since model *H* has the wrong unit cell and wrong symmetry, and model *G* the wrong unit cell and wrong atom content, these two are in fact the most flawed of all the proposed structural models!

There is an additional point of interest concerning the disregarding of systematically weak reflections when using the models based on the smaller unit cell. When using the correct

unit-cell choice, the ratio of observed to measured reflections is quite low (0.25 and 0.20 for data measured by Mo *K*α and Cu *K*α radiation, respectively). However, when a structural model based on the smaller unit cell is used, the ratio of observed to measured reflections increases. The increase is only slight for Mo *K*α data (0.28), but for Cu *K*α data it is much more pronounced (0.34). In the *checkCIF/PLATON* structure validation suite a ratio of observed to measured reflections below 0.30 will generate an ALERT of the level A, while one above 0.3 generates only a level B ALERT. Therefore, if one only looks at the *checkCIF/PLATON* report, without bearing in mind the entirely artificial genesis of the ‘improvement’ of the data, one might conclude that the models *B*, *G* and *H* refined against Cu *K*α radiation data are based on substantially higher-quality data than the correct model.

4. Conclusion

Although numbers are proverbially not supposed to lie, it would seem that in this case they are not to be entirely trusted. Refinement of wrong structural models may lead to better statistical parameters (*R*, *wR*, *S*) and thus to a misconception on the true nature of a crystal structure. Although we have demonstrated this on a crystal designed with the intention to make the (correct) structure determination as difficult as possible, it is conceivable that similar effects will be encountered in many cases, in particular when dealing with weakly diffracting crystals and crystals with periodical superstructures which lead to systematically weak reflections, disregarding which may lead to an apparent increase of measured data quality. This effect is present even when using a high-intensity source (such as microfocus Cu radiation), although such measurement gives superior quality data. If, in such a case, the gathered diffraction data are examined with great scrutiny, and if the structure solution and refinement are approached with due consideration of the chemical and physical properties of the matter in question, rather than a mere glance at the value of *R*, the probability of reaching an incorrect structural model should be substantially reduced.

The authors would like to thank Krešimir Molčanov of the Ruder Bošković Institute in Zagreb for performing the measurements using Cu radiation, and the Ministry of Science, Education and Sport, Republic of Croatia, for financial support of this work through grant No. 119-1193079-3069.

References

- Allen, F. H. (2002). *Acta Cryst.* **B58**, 380–388.
- Clemente, D. A. & Marzotto, A. (2004). *Acta Cryst.* **B60**, 287–292.
- Farrugia, L. J. (1997). *J. Appl. Cryst.* **30**, 565.
- Herbstein, F. H. & Marsh, R. E. (1998). *Acta Cryst.* **B54**, 677–686.
- Lloyd, M. A. & Brock, C. P. (1997). *Acta Cryst.* **B53**, 773–779.
- Marsh, R. E. (1995). *Acta Cryst.* **B51**, 897–907.
- Marsh, R. E. (2002). *Acta Cryst.* **B58**, 893–899.
- Marsh, R. E. & Herbstein, F. H. (1988). *Acta Cryst.* **B44**, 77–88.

- Marsh, R. E., Kapon, M., Hu, S. & Herbstein, F. H. (2002). *Acta Cryst.* **B58**, 62–77.
- Oxford Diffraction (2003). *CrysAlis CCD and CrysAlis RED*. Version 1.170. Oxford Diffraction Ltd, Wroclaw, Poland.
- Persistence of Vision Pty (2004). *Persistence of Vision Raytracer (POV-Ray)*. Version 3.6. <http://www.povray.org/download/>.
- Schnering, H. G. von & Vu, D. (1983). *Angew. Chem.* **95**, 421.
- Sheldrick, G. M. (1997). *SHELXS97* and *SHELXL97*. University of Göttingen, Germany.
- Sheldrick, G. M. (2008). *Acta Cryst.* **A64**, 112–122.
- Spek, A. L. (1998). *PLATON98: a Multipurpose Crystallographic Tool*. Version 120398. University of Utrecht, Utrecht, The Netherlands.
- Spek, A. L. (2003). *J. Appl. Cryst.* **36**, 7–13.
- Spek, A. L. (2009). *Acta Cryst.* **D65**, 148–155.