STRUCTURAL CHEMISTRY
ISSN 2053-2296

Crystallographic curiosities: polymorphism and structures with $Z' > 1$

William Clegg*

School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK.
*Correspondence e-mail: bill.clegg@ncl.ac.uk

Keywords: polymorphism; crystal structure; CSD; CCDC; $Z' > 1$; zprime.

Crystallography cuts across a wide range of scientific disciplines, encompassing aspects of biology and medicine, chemistry, physics, earth and environmental sciences, materials science, and engineering. A major focus within chemistry is the determination of crystal structures from diffraction data, and a significant milestone was reached this June when the one millionth structure was added to the Cambridge Structural Database (Groom *et al.*, 2016; https://www.ccdc.cam.ac.uk/, accessed on 20 June 2019). This rapidly expanding resource offers not only a convenient single-source collection of all individual published and deposited organic and metal–organic crystal structures, but also a flexibly searchable information-rich mine of data for exploring patterns, trends and relationships among families of structures and aspects of their properties. It has been extensively used for studies of molecular conformation, analyses of intermolecular interactions, modelling of some kinds of chemical reaction pathways, and the generation of average and idealized geometries for structure prediction and energy calculations, among other applications. Its detailed contents are amenable to a whole gamut of statistical analyses, with the caveat that it does not necessarily provide an unbiased sample of all crystal structures, nor even of all that have been experimentally determined, given that a high proportion of such results never find their way into publications and those that do have effectively been selected for their interest and importance to authors, publishers and readers of journals.

Examination of this huge library of structural results throws up a number of interesting questions and curiosities. Among these are the topics of polymorphism – the occurrence of more than one crystal structure for a given chemical compound – and the incidence of structures with more than one molecule, unrelated by crystallographic symmetry, in the asymmetric unit ($Z' > 1$). Each of these is far more than just an intellectual curiosity and a subject for heated debate on matters of definition, scope and terminology; they are of considerable practical importance. Pharmaceutical companies, in particular, expend much time, energy and money in what they hope are exhaustive polymorph screening tests in order to ensure that their commercial products have the most appropriate physical properties and are fully covered by patents. Structures with $Z' > 1$ have a major impact on enterprises such as crystal structure prediction and the *ab initio* solution of crystal structures from powder diffraction data. The two topics are not unrelated, as polymorphs are often found to have different values of Z'.

They feature together in a report in the previous issue of *Acta Crystallographica Section C* by Alvarenga *et al.* (2019) exploring some structural aspects of a thiophene-substituted benzothiazole with potential biomedical applications (Fig. 1). This also includes the topic of cocrystallization, but that is another saga extending beyond the space limitations of this commentary! Three new polymorphs are reported, having Z' values of 8, 4 and 4, and are compared with a previously known $Z' = 2$ polymorph [Huang *et al.* (2016); also reported as a *CSD Communication* by Renz *et al.* (2011)]. All these polymorphs have the molecule in the same conformation, an example of so-called packing polymorphism in contrast to the more commonly observed polymorphism of flexible molecules displaying different conformations. Contributions to the lattice energies of the structures from intermolecular interactions are assessed through theoretical calculations. In the absence of classical hydrogen bonding, these are all of relatively weak secondary types: face-to-face π–π stacking, C—H$\cdots\pi$ and S$\cdots\pi$ involving aromatic rings, C—H$\cdots$S and S$\cdots$S atom-to-atom interactions. Each of these contributes less than 10 kcal mol^{-1} to the overall stability of the crystal structure, π–π being the strongest with the simultaneous involvement of all three rings in each molecule. The plethora of

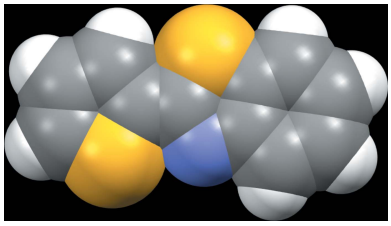

Table 1
Organic structures with $Z' > 1$ in the CSD.

Z'	CSD entries (organic)	Z'	CSD entries (organic)
2	49369	10	15
3	3051	11	2
4	3689	12	32
5	147	13	0
6	358	14	3
7	16	15	2
8	200	16	13
9	17	>16	8

available intermolecular interactions covering only a small range of energies is seen as a factor in the occurrence of several polymorphs. It is also probably a reason for the disorder observed in all of the new polymorphs, with two essentially parallel orientations of the molecule in each of its locations. (This is treated as a whole-molecule disorder, with a *transoid* arrangement of the two S atoms around the bond linking the two five-membered rings in all cases, and no contribution of a *cisoid* conformation; a search of the CSD for related structures with two S-containing five-membered rings linked by a bond not itself part of any ring shows that *transoid* conformations with an S—C—C—S torsion angle within ±25° of 180° are over 85% of the total, and almost all others have two thiophene rings, validating this model.)

The results and discussion provided by Alvarenga *et al.* (2019) are a significant contribution to the ongoing debates about polymorphism and the incidence of $Z' > 1$ structures. While many scientists prefer to restrict the term polymorphism to families of structures of the same chemical compound having the same overall composition for the structure as a whole, others (including the United States Food and Drugs Administration) broaden this to cover solvates, sometimes referred to as pseudopolymorphs; a comprehensive discussion and systematic review is provided by Brog *et al.* (2013). Summary statistics provided on the CSD website give a total of 10870 polymorph families represented in the database; these are polymorphs in the narrower sense, solvates being classified separately. Attempts to understand some of the aspects of polymorphism have included previous analyses of the contributions of intermolecular interactions to the total lattice energy and examination in particular of polymorphs with $Z' > 1$ (Bernstein *et al.*, 2008; Steed & Steed, 2015). The Z' story itself is a long and fascinating one, with various explorations of the energetic, kinetic and other factors involved (Anderson *et al.*, 2011; Taylor *et al.*, 2016), and it enjoys its own dedicated website (J. W. Steed, http://zprime.co.uk/, accessed on 20 June 2019). One example of the quirks of Z' is the distribution of values greater than 1; Table 1 shows this for organic entries in the CSD (November 2018 version with updates to May 2019), in which there are relatively high values for $Z' = 8$, 12 and 16, and low values for odd numbers except 3 and perhaps 9. Some of the factors that have been identified or suggested as contributors to an increased probability of a structure having $Z' > 1$ include awkward molecular shape, frustrated preferences for relatively strong intermolecular interactions, such as classical hydrogen bonds, pseudo-inversion arrangements for a molecule containing one chiral centre, and an abundance of alternative weak intermolecular interactions (as in the polymorphs described here). The incidence of $Z' > 1$ seems to have some correlation with other problematic aspects of crystal structure determinations, such as small crystal size, poor crystal quality, disorder and the need for synchrotron radiation (Nichol & Clegg, 2007). Since the time of that simple survey, chemical single-crystal diffraction synchrotron facilities have become more widely available and the number of synchrotron-derived entries in the CSD has grown more than tenfold, while the size of the CSD as a whole has increased by a factor of 2.5. Although the exact proportions have shifted somewhat, it remains the case that structures from synchrotron data have a significantly higher proportion with $Z' > 1$ than those from laboratory X-ray sources, while neutron-derived structures have a significantly lower proportion. With a much larger total population for these statistics, there is greater confidence in the suggestion that $Z' > 1$ may often be a marker of 'difficult' nonroutine structures, such as these new thiophenyl–benzothiazole polymorphs.

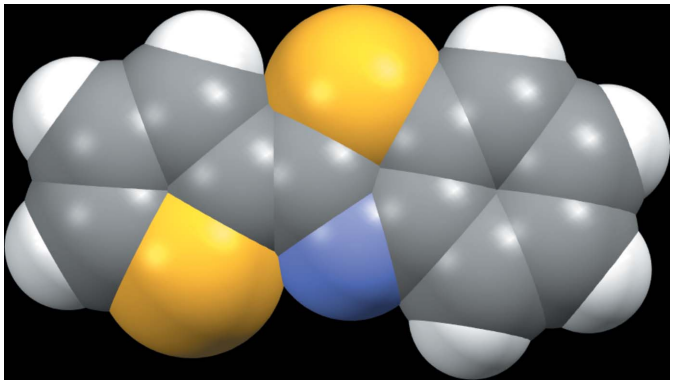

Figure 1
Space-filling representation of a molecule of 2-(thiophen-2-yl)-1,3-benzothiazole.

References

Alvarenga, M. E., Valdo, A. K. S. M., Ribeiro, L., Do Nascimento Neto, J. A., De Araujo, D. P., Da Silva, C. M., De Fátima, Â. & Martins, F. T. (2019). *Acta Cryst.* C**75**, 667–677.

Anderson, K. M., Probert, M. R., Goeta, A. E. & Steed, J. W. (2011). *CrystEngComm*, **13**, 83–87.

Bernstein, J., Dunitz, J. D. & Gavezzotti, A. (2008). *Cryst. Growth Des.* **8**, 2011–2018.

Brog, J.-P., Chanez, C.-L., Crochet, A. & Fromm, K. M. (2013). *RSC Adv.* **3**, 16905–16931.

Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. (2016). *Acta Cryst.* B**72**, 171–179.

Huang, Y.-T., Wang, W.-C., Hsu, C.-P., Lu, W.-Y., Chuang, W.-J., Chiang, M. Y., Lai, Y.-C. & Chen, H.-Y. (2016). *Polym. Chem.* **7**, 4367–4377.

Nichol, G. S. & Clegg, W. (2007). *CrystEngComm*, **9**, 959–960.

Renz, F., Baran, Y., Gutlich, P. & Schollmeyer, D. (2011). *CSD Communication* (refcode CAMBAV; CCDC deposition number 815737). CCDC, Cambridge, England.

Steed, K. M. & Steed, J. W. (2015). *Chem. Rev.* **115**, 2895–2933.

Taylor, R., Cole, J. C. & Groom, C. R. (2016). *Cryst. Growth Des.* **16**, 2988–3001.