EuroSciCon

3rd Edition of International Conference on

Advanced Spectroscopy, Crystallography and Applications in Modern Chemistry

June 04-05, 2018
London, UK

Jan Rohlicek, Struct Chem Crystallogr Commun 2018, Volume 4
DOI: 10.21767/2470-9905-C1-005

CRYSTALCMP – FAST PACKING COMPARISON OF MOLECULAR CRYSTALS

Jan Rohlicek
Institute of Physics, ASCR, Czech Republic

The crystal structure similarity is not an established and generally defined property. There are various definitions of crystal structure similarity defined for different purposes, each having different advantages and disadvantages in different situations. There are several methods, that define the similarity of crystal structures as a similarity of a representative functions called fingerprints [1]–[4]. These methods compare crystal structures indirectly, by comparison of their fingerprints. Other methods, are trying to compare atomic coordinates [5]–[8] or even positions of basic moieties in the crystal structures [9], [10]. In all cases, when differences in positions of atoms or moieties are used for calculation of the crystal structure similarity, the transformation between crystal structures has to be determined. The difficulty of this procedure is nicely described in [8]. CrystalCMP [11] is a software for comparison of molecular packing that was recently published. The suggested method is based on the second mentioned approach - comparison of molecular positions. It is immediately clear, that the comparison method is designed for all non-polymeric crystal structures, where some stand-alone moieties (molecular fragments) can be found. It is perfectly valid for all molecular crystals and some of the metal-organic complexes. Most of the inorganic structures and MOF with polymeric structures cannot be compared by this method. The comparison method is divided in several steps: (i) Definition of the central molecule (the largest molecule in the unit cell by default), (ii) creating of the molecular cluster (10 surrounding molecules by default), which is representing the whole crystal structure, (iii) definition of the fragment for overlaying (either by SMILES notation or by HASH strings as originally published in 2016) and (iv) overlapping molecular clusters according to the defined fragment and (v) calculating differences in molecular positions and its relative rotations, see definition of the Psab formula.

$$Ps_{a,b} = D_c + X \cdot \frac{A_d}{180}$$

where Dc is the average distance (in Å) between the molecular centers of related molecular pairs and Ad is the average angle (in degrees) between them. The X value is set by the user to weight the influence of the Ad parameter (the default value is X = 100), see Fig 1. As a result of comparison is a similarity matrix with calculated dendrogram and the transformation matrix between both compared molecular clusters. This enables overlaying the compared structures and see differences visually in human-readable form. The advantage of this method is its low sensitivity to the relatively large expansion of the molecular structure caused e.g. by the temperature or even by the presence of different solvent molecules in the crystal structure. For that reason this method is applicable for comparison of solvatomorphic series of identical or even just similar compounds. Several tests on different compounds had been performed. The algorithm compares two molecular packing in less than one second on a common office PC (approx. 100 ms for small molecule of benzamide and approx. 200 ms for middle-size molecule of trospium [11]). This allows making comparison of large number of compounds. In addition, automation of the method allows, for example, comparison of all crystal structures in the whole CSD database

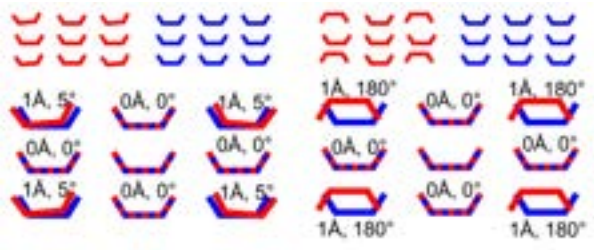

Figure 1 A graphical interpretation of the method used in CrystalCMP. (Top left and top right) In both cases the central molecules are surrounded by eight molecules. (Bottom) Both upper clusters are overlapped and the numbers near each surrounded molecule represent the differences in the molecular centers of related pairs and the angle differences between them. In general, these numbers are real. (Left) Two crystal structures with almost identical packing, Psab = 0.5 + 100 × (2.5/180) = 1.9 (for X = 100). (Right) Two crystal structures with almost identical positions of surrounding molecules, but with different packing of surrounding molecules, Psab = 0.5 + 100 × (90/180) = 50.5 (for X = 100).

Recent Publications

1. M. Valle and A. R. Oganov, "Crystal fingerprint space – a novel paradigm for studying crystal-structure sets," Acta Crystallogr. A, vol. 66, no. 5, pp. 507–517, Sep. 2010.
2. E. L. Willighagen, R. Wehrens, P. Verwer, R. de Gelder, and L. M. C. Buydens, "Method for the computational

comparison of crystal structures," Acta Crystallogr. B, vol. 61, no. 1, pp. 29–36, Feb. 2005.

3. H. R. Karfunkel, B. Rohde, F. J. J. Leusen, R. J. Gdanitz, and G. Rihs, "Continuous similarity measure between nonoverlapping X-ray powder diagrams of different crystal modifications," J. Comput. Chem., vol. 14, no. 10, pp. 1125–1135, Oct. 1993.
4. R. de Gelder, R. Wehrens, and J. A. Hageman, "A generalized expression for the similarity of spectra: application to powder diffraction pattern classification," J. Comput. Chem., vol. 22, no. 3, pp. 273–289, Feb. 2001.
5. A. V. Dzyabchenko, "Method of crystal-structure similarity searching," Acta Crystallogr. B, vol. 50, no. 4, pp. 414–425, Aug. 1994.
6. R. Hundt, J. C. Schön, and M. Jansen, "CMPZ – an algorithm for the efficient comparison of periodic structures," J. Appl. Crystallogr., vol. 39, no. 1, pp. 6–16, Feb. 2006.
7. B. P. Van Eijck and J. Kroon, "Fast clustering of equivalent structures in crystal structure prediction," J. Comput. Chem., vol. 18, no. 8, pp. 1036–1042, Jun. 1997.
8. G. de la Flor, D. Orobengoa, E. Tasci, J. M. Perez-Mato, and M. I. Aroyo, "Comparison of structures applying the tools available at the Bilbao Crystallographic Server," J. Appl. Crystallogr., vol. 49, no. 2, pp. 653–664, Apr. 2016.
9. C. F. Macrae et al., "Mercury CSD 2.0 – new features for the visualization and investigation of crystal structures," J. Appl. Crystallogr., vol. 41, no. 2, pp. 466–470, Apr. 2008.
10. J. A. Chisholm and S. Motherwell, "COMPACK : a program for identifying crystal structure similarity using distances," J. Appl. Crystallogr., vol. 38, no. 1, pp. 228–231, Feb. 2005.
11. J. Rohlíček, E. Skořepová, M. Babor, and J. Čejka, "CrystalCMP : an easy-to-use tool for fast comparison of molecular packing," J. Appl. Crystallogr., vol. 49, no. 6, pp. 2172–2183, Dec. 2016.

Biography

Jan Rohlicek has his expertise in crystal structure determination from powder diffraction data. He is the author of a grid extension of program FOX, that is used for crystal structure determination from powders. He is also author of the program MCE (Marching Cule ELD) for placing atoms and fragments to the 3D Fourier maps and of the presented program CrystalCMP for comparison of molecular packing. He is responsible for the laboratory of powder diffraction at the Department of Structure Analysis at the Institute of Physics ASCR in Prague

rohlicek@fzu.cz