xINTERPDF: a GUI program for analyzing intermolecular pair distribution functions in organic compounds from X-ray total scattering data

Chenyang Shi*

Drug Product Development, Research and Development, AbbVie Inc., North Waukegan Road, North Chicago, IL 60064, USA. *Correspondence e-mail: chenyang.shi@abbvie.com

1. The crystallographic problem

Structures of organic compounds are more complex than their inorganic counterparts, which have usually a network structure, representing a giant "molecule". Organics, on the other hand, have strong intramolecular bonds but much weaker intermolecular interactions, making them prone to structural disorders. Another complexity comes from the weak X-ray scattering of light elements (C, H, O, N etc) which are the building blocks of organic compounds. The atomic pair distribution function (PDF) calculated from X-ray total scattering has been demonstrated to be a valuable tool for investigating structures of disordered and amorphous organics compounds (Shi *et al.*, 2017; Prill *et al.*, 2015; Prill *et al.*, 2016). Although existing tools such as DiffPy-CMI (Juhás *et al.*, 2016) and XISF (Mou *et al.*, 2015) can be used for this problem, a new software program is still of great value that provides a user-friendly graphical user interface (GUI, as opposed to command-driven in DiffPy-CMI) and analyzes the data in real-space (as opposed to reciprocal space in XISF).

2. Method of solution

In the program xINTERPDF, a user-friendly graphical user interface has been built to facilitate user interactions with the data. It currently supports the study of intermolecular interaction (e.g. hydrogen bonds) by subtracting out the simulated PDF signal of a single molecule from the measured total PDF. The PDFs of given structures can be calculated in either real space or reciprocal space. The intermolecular PDFs are updated and visualized in real time when users fine tune the parameters. Additionally, it provides a straightforward interface for performing a PDF model fit to the X-ray PDF data for the crystalline organic compounds (Prill *et al.*, 2015). In both cases, the plots and associated raw data can be saved easily. The program is written in open source Python programming language, and is easily distributed to various systems using conda (Conda, 2018).

3. Software and hardware environment

The program runs on both Linux and macOS computers. It is written in Python 2.7, using its default Tkinter module to create the GUI, matplotlib for visualization, NumPy and SciPy for scientific calculations (Python, 2018). It heavily relies on DiffPy-CMI package (Juhás *et al.*, 2016) for the simulations of PDFs.

4. Program specification

xINTERPDF should run in the same way on Linux and macOS systems. The look and feel of the GUI may slightly vary. When studying intermolecular PDF, the simulation of PDF in real space is finished instantaneously. However, it takes much longer time for simulating PDF of a crystal using Debye scattering equation (Debye, 1915) in Q-space. For a typical model fit of crystalline PDF (e.g. 104 atoms

in the expanded cell) in a fit range up to 40 Å, it takes about ~10 minutes to complete on macOS 10.10.3 and 64 bit system with 3.1 GHz Intel Core i7 and 16 GB memory.

5. Documentation and availability

The xINTERPDF program has a homepage at https://github.com/curieshicy/xINTERPDF. Details about installation and applications are explained in its User Guide (https://github.com/curieshicy/xINTERPDF/tree/master/User_Guide_0.1.0).

Disclosure

C.S. is the employee of AbbVie and may own AbbVie stock. The design, study conduct, and financial support for this research were provided by AbbVie. AbbVie participated in the interpretation of data, review, and approval of the publication.

Acknowledgements

The author would like to thank 11-ID-B beamline staff Olaf Borkiewicz, Kevin Beyer and Karena Chapman at Argonne National Laboratory for their assistance in carrying out the PDF experiments. Dr. Pavol Juhás, Prof. Lian Yu and Prof. Kirsten Jensen are thanked for their valuable suggestions on improving the program. Use of the Advanced Photon Source, an Office of Science User Facility operated for the US Department of Energy (DOE) Office of Science by Argonne National Laboratory, was supported by the US DOE under contract No. DE-AC02-06CH11357.

References

Shi, C., Teerakapibal, R., Yu, L. & Zhang, G. G. Z. (2017). *IUCrJ* **4**, 555-559.

Prill, D., Juhás, P., Schmidt, M. U. & Billinge, S. J. L. (2015). *J. Appl. Cryst.* **48**, 171-178.

Prill, D., Juhás, P., Billinge, S. J. L. & Schmidt, M. U. (2016). *Acta Crystallogr. A* **72**, 62-72.

Juhás, P., Farrow, C. L., Yang, X., Knox, K. R. & Billinge, S. J. L. (2015). *Acta Crystallogr. A* **71**, 562-568.

Mou, Q., Benmore, C. J. & Yarger J. L. (2015). *J. Appl. Cryst.* **48**, 950-952.

Conda (2018). https://conda.io/docs/.

Python (2018). (a) https://www.python.org/ (b) Tkinter https://docs.python.org/2/library/tkinter.html (c) matplotlib https://matplotlib.org/ (d) numpy http://www.numpy.org/ (e) Scipy https://www.scipy.org/.

Debye, P. (1915) **351**, 809-823.