



# *xINTERPDF*: a graphical user interface for analyzing intermolecular pair distribution functions of organic compounds from X-ray total scattering data

Chenyang Shi\*

Drug Product Development, AbbVie Inc., 1 North Waukegan Road, North Chicago, IL 60064, USA. \*Correspondence e-mail: [chenyang.shi@abbvie.com](mailto:chenyang.shi@abbvie.com)

Received 9 March 2018

Accepted 30 August 2018

Edited by Th. Proffen, Oak Ridge National Laboratory, USA

**Keywords:** X-ray total scattering; pair distribution function; Python programming; graphical user interfaces.

A new software program, *xINTERPDF*, that analyzes the intermolecular correlations in organic compounds *via* measured X-ray total scattering data is described.

## 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 disorder. 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 (and/or neutron, electron) total scattering has been demonstrated to be a valuable tool for investigating structures of disordered and amorphous organic compounds (Shi *et al.*, 2017; Prill *et al.*, 2015, 2016; Benmore & Weber, 2011; Rademacher *et al.*, 2012; Chen *et al.*, 2014; Gorelik *et al.*, 2015). Although existing tools such as *DiffPy-CMI* (Juhás *et al.*, 2015) and *XISF* (Mou *et al.*, 2015) can be used for this problem, a new software program that provides a user-friendly graphical user interface (GUI, as opposed to script-based programming in *DiffPy-CMI*) and analyzes the data in real space (as opposed to in reciprocal space in *XISF*) is still of great value. This article describes such a program, *xINTERPDF*.

## 2. Method of solution

In *xINTERPDF*, user-friendly GUIs have been built to facilitate user interactions with the data. It currently supports the following:

- (1) The study of intermolecular interaction (*e.g.* hydrogen bonds) by subtracting out the scattering signal of a single molecule in real space.
- (2) The PDF model fit of the crystalline organic compound using the method proposed by Prill *et al.* (2015).
- (3) The phase quantification of physical mixtures of organics.
- (4) The generation of score/scree plots based on principle component analysis (PCA).

The program is written in the open-source Python programming language (<https://www.python.org/>) and is



OPEN ACCESS

distributed to various operating systems using the *Conda* package manager (<https://conda.io/docs/>).

### 3. Software and hardware environment

The program runs on both 64-bit Linux and macOS machines. It is written in Python 2.7, using its default *Tkinter* module (<https://docs.python.org/2/library/tkinter.html>) to create the GUI, *Matplotlib* (<https://matplotlib.org/>) for visualization, and *NumPy* (<http://www.numpy.org/>) and *SciPy* (<https://www.scipy.org/>) for scientific calculations. The `sklearn.decomposition.PCA` module from *Scikit-Learn* (Pedregosa *et al.*, 2011) is called for application of PCA. The *DiffPy-CMI* package (Juhás *et al.*, 2015) is used as a backend for the simulations of PDFs.

### 4. Program specification

*xINTERPDF* runs in the same way on Linux and macOS systems. The look and feel of the GUI may slightly vary. When studying intermolecular interactions in organics, the simulation of the PDF in real space is finished almost instantaneously. However, it takes a relatively longer time for simulating the PDF of a crystal using the Debye scattering equation (Debye, 1915). For a typical model fit of a crystalline PDF (e.g. D-mannitol with 104 atoms in the expanded cell) in an  $r$  range up to 40 Å, it takes about ~10 min to complete on macOS 10.10.3 with a 3.1 GHz Intel Core i7 and 16 GB memory. The usages of phase quantification and PCA return results in real time.

### 5. Documentation and availability

The home page for the *xINTERPDF* program is <https://www.diffpy.org/products/xinterpdf.html>, where users may find instructions for installation and the manual for applications. The source code is hosted at GitHub page <https://github.com/curieshicy/xINTERPDF>.

### 6. Disclosure

C. Shi 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, Professor Lian Yu and Dr Fabia Gozzo are thanked for their valuable suggestions on improving the program.

### Funding information

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

- Benmore, C. J. & Weber, J. K. R. (2011). *Phys. Rev. X*, **1**, 011004.  
 Chen, S., Sheikh, A. Y. & Ho, R. (2014). *J. Pharm. Sci.* **103**, 3879–3890.  
 Debye, P. (1915). *Ann. Phys.* **351**, 809–823.  
 Gorelik, T. E., Schmidt, M. U., Kolb, U. & Billinge, S. J. L. (2015). *Microsc. Microanal.* **21**, 459–471.  
 Juhás, P., Farrow, C., Yang, X., Knox, K. & Billinge, S. (2015). *Acta Cryst.* **A71**, 562–568.  
 Mou, Q., Benmore, C. J. & Yarger, J. L. (2015). *J. Appl. Cryst.* **48**, 950–952.  
 Pedregosa, F. *et al.* (2011). *J. Mach. Learn. Res.* **12**, 2825–2830.  
 Prill, D., Juhás, P., Billinge, S. J. L. & Schmidt, M. U. (2016). *Acta Cryst.* **A72**, 62–72.  
 Prill, D., Juhás, P., Schmidt, M. U. & Billinge, S. J. L. (2015). *J. Appl. Cryst.* **48**, 171–178.  
 Rademacher, N., Daemen, L. L., Chronister, E. L. & Proffen, Th. (2012). *J. Appl. Cryst.* **45**, 482–488.  
 Shi, C., Teerakapibal, R., Yu, L. & Zhang, G. G. Z. (2017). *IUCrJ*, **4**, 555–559.