TECHNICAL ARTICLE

Genie: new software for rapid submission of powder diffraction data to the ICDD powder diffraction file™

Diane E. Sagnella,[a] and Emily Foster
International Centre for Diffraction Data, Newtown Square, Pennsylvania

(Received 17 June 2014; accepted 8 September 2014)

A new software application has been developed to provide a convenient method for the submission of diffraction data directly to the International Centre for Diffraction Data, for inclusion in the Powder Diffraction File (PDF)™. The application, titled Genie, acts as a portal for the merging of data from separate sources allowing the user to quickly and easily submit their data for the PDF. Using Java Web Start technology Genie can be run using all common computer platforms. Genie is flexible and can be used to read a variety of common file formats. Furthermore, Genie can also be used for submission of New Diffraction Data for publication in the *Powder Diffraction Journal*. © *2014 International Centre for Diffraction Data*. [doi:10.1017/S0885715614000864]

Key words: PDF, Powder Diffraction, Java Web Start

I. INTRODUCTION

The Powder Diffraction File (PDF) has been the primary reference for quality powder diffraction data for over 70 years (Hanawalt, 1983). It is used by the scientists worldwide to identify phases in their samples. Initially, the PDF was a collection of single-phase X-ray powder diffraction patterns consisting of a reduced data set of *d*-spacings and intensities.

Over the past years the major component of the PDF were experimental patterns. To this day, significant effort is still put toward the publication of editorially reviewed and high-quality experimental patterns. However, over time fewer diffraction pattern reflection lists are published in quality journals. Not long ago, the focus of much research was the indexing of patterns and it was common to find complete reflection lists published in the papers. This allowed the data from these published X-ray powder diffraction experiments, including the diffraction peak reflection list, unit cells, etc. to be conveniently abstracted. This leads to approximately 2000 *new* experimental patterns being added to the PDF each year.

However, with the advent of techniques such as Rietveld refinement (Rietveld, 1969), the focus of the scientific community has shifted from indexing a compound's diffraction pattern to determining its crystal structure. As a result, the number of experimental powder patterns published each year has been on the decline. For example, in 2003 almost 2500 experimental X-ray powder diffraction patterns were added to the PDF, whereas only 1513 were added in 2013 (ICDD, 2013).

As the needs of the scientific community and the available analytical technologies have changed, the International Centre for Diffraction Data (ICDD) has continued to update the PDF with new and improved information. To better enable users to perform analyses such as the Rietveld method, starting in 2005, the ICDD began to include atomic coordinates in some PDF entries.

In fact, with the 2014 release of the PDF4 + , the ICDD has incorporated over 190 000 inorganic and 59 000 organic crystal structures as complementary data to the corresponding powder diffraction data in the database (ICDD, 2014). Their inclusion in the PDF4 line of products has provided the user with the tools necessary for quantitative phase analysis using the Rietveld method (Bish and Howard, 1988) and whole pattern fitting. However, as useful as atomic coordinates are, they do not fulfill the primary purpose of the PDF: the phase identification of materials which is best done through comparison of powder diffraction results to experimental data.

Therefore as the number of atomic coordinates included in the PDF continues to climb, we must continue to find and include experimental X-ray diffraction data. With fewer authors publishing their diffraction data, how can we do this?

One effort to address this issue has been to perform exhaustive bibliographic searches to find papers that mention X-ray diffraction experiments on systems of commercial importance but where data are unpublished. The ICDD then contacts the authors of these papers and requests material not found in the publication, such as the reflection list, the powder data and other supplementary information. This process is labor intensive and the chance of receiving a response from the author is strongly dependent on the time lag between publication of the paper and the date the request was sent.

Therefore careful consideration has gone into the development of an improved method of getting high-quality *experimental* diffraction data into the PDF. Clearly, the scientific community needs a fast and user-friendly method of submitting data directly to the ICDD, either before or after its publication. Ideally, the method would be always ready for the user, available anytime and anywhere. The subject of this paper is an online access program, Genie, which allows users to conveniently submit their experimental diffraction data, and optionally crystal structure fits with the possibility of publishing their results as a *Rapid Communication* in this journal. Herein we

[a]Author to whom correspondence should be addressed. Electronic mail: sagnella@icdd.com

TABLE I. File formats currently supported by Genie.

All data	Powder data	Reflection list	Coordinates
CIF	CIF	CIF	CIF
PDB	CPI	*HKL* (FullProf)	PDB
	GSAS	TXT	TXT
	TXT		

describe the Genie program and how it is used to submit diffraction data for review by ICDD's editors for possible inclusion in the PDF *and* publication in *Powder Diffraction.* Since publications are both the method for archival of research results as well as an important metric of productivity, it is hoped that this will encourage scientists to communicate their results in this manner.

A. Genie

Genie is a graphical user interface (GUI) program developed by the ICDD that allows a researcher to quickly and easily submit data for publication in the PDF. Genie can also be used to submit data for publication in the *Powder Diffraction Journal*, thus providing the user with a quick and efficient method for publishing their data as a rapid communication. In fact, using Genie, the user can see their work published in as little as 3 months.

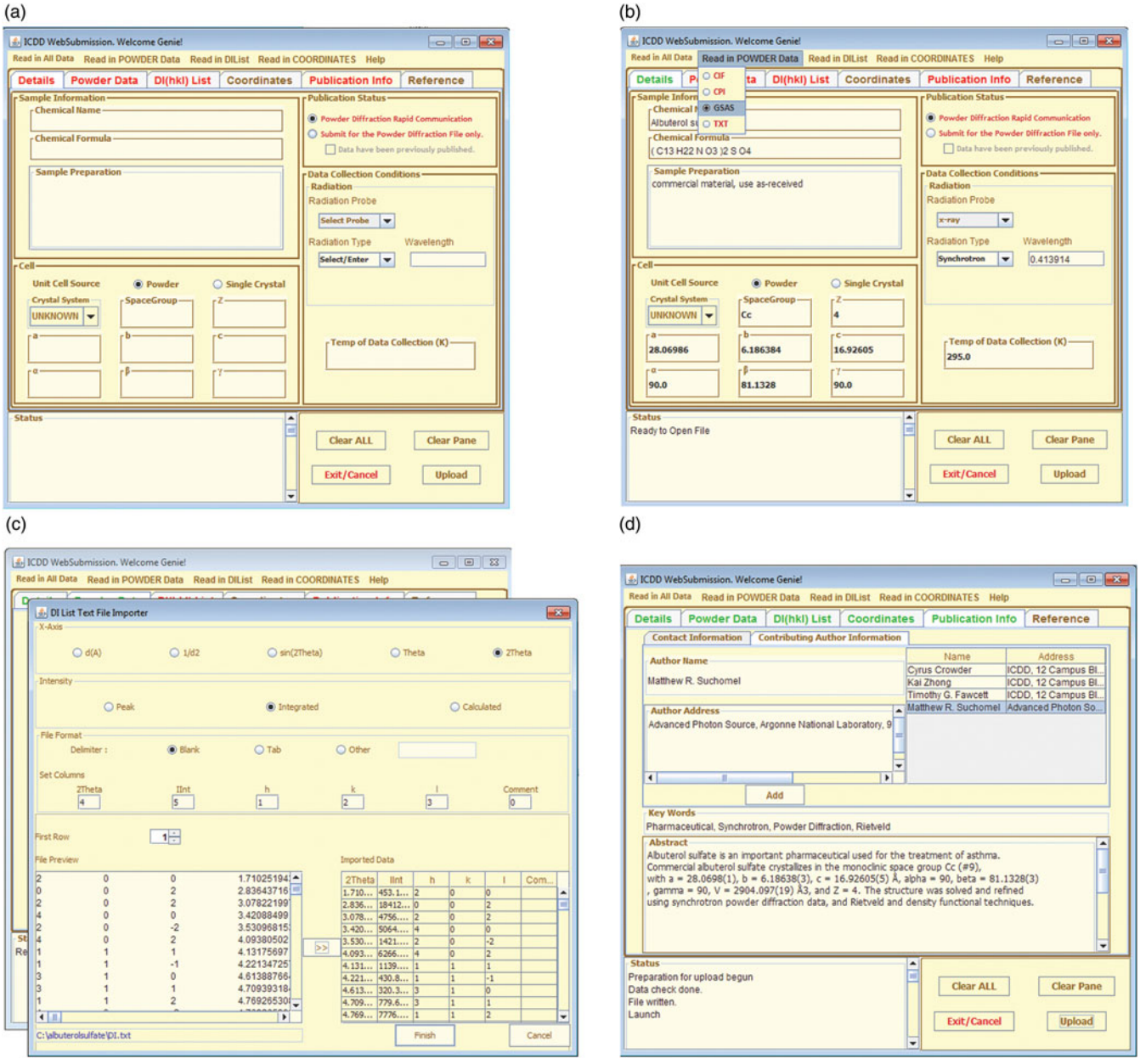

Figure 1. (Color online) (a–d) The Genie Interface. (a) The main pane presented to the user. In this instance, the user has indicated the submission be considered from publication in *Powder Diffraction Journal.* As a result, the tabs that require data are highlighted in red. (b) Having manually entered data in the details pane, the user moves on to read in powder data from a GSAS file. (c) The reflection list is being read in from a text file using the Genie text file importer. (d) The coordinate pane is highlighted in green indicating that they have been read in as well. While coordinates are not *required* for a rapid communication, their inclusion enhances the submission. For submission to the *Powder Diffraction Journal*, the user enters contact information and an abstract for the rapid communication.

https://doi.org/10.1017/S0885715614000864 Published online by Cambridge University Press

Genie, a Java Web Start application, can be run anywhere (Genie has been tested on Windows XP, Windows Vista, Windows 7, Windows 8, Mac OS X Mountain Lion, Snow Leopard & Maverick and Oracle Linux.). The choice of Java Web Start offers several distinct advantages over other programming languages. Java Web Start technology provides a framework that allows users to download and run full Java™ programs directly from the internet using a web browser. This means that the Genie program can be launched with a simple button click from the user's web browser. Note that Java programs are compiled into a *platform-independent* binary format, all that is needed is a Java client, known as Java Virtual Machine (JVM)(Java Runtime Environment (JRE) 7 or higher is required.) for the local computer. These are very widely available for computers and even more modern computing platforms such as tablets and phones. This means

RAPID COMMUNICATION

Powder X-ray diffraction of albuterol sulfate $(C_{13}H_{22}NO_3)_2SO_4$

J.A. Kaduk,[1] C.E. Crowder,[2] K. Zhong,[2,a)] T.G. Fawcett,[2] and M.R. Suchomel[3]
[1]*Illinois Institute of Technology, 3101 S. Dearborn Street, Chicago, Illinois 60616*
[2]*ICDD, 12 Campus Boulevard, Newtown Square, Pennsylvania 19073-3273*
[3]*Argonne National Laboratory, Advanced Photon Source, 9700 S. Cass Avenue, Argonne, Illinois 60439*

(Received 10 March 2014; accepted 12 March 2014)

Albuterol sulfate (marketed as ProAir HFA or Ventolin HFA) is an important pharmaceutical used for the treatment of asthma. Commercial albuterol sulfate crystallizes in the monoclinic space group Cc (9), with $a = 28.0698(1)$ Å, $b = 6.18638(3)$ Å, $c = 16.92605(5)$ Å, $\alpha = 90°$, $\beta = 81.1328(3)°$, $\gamma = 90°$, $V = 2904.097(19)$ Å^3, and $Z = 4$. Previous structure determination did not include hydrogens (Leger et al., 1978). In this work, the sample was ordered from Sigma-Aldrich. The structure was solved and refined using synchrotron ($\lambda = 0.413914$ Å) powder diffraction data, and Rietveld and density functional techniques. Figure 1 shows the Powder X-ray diffraction pattern of the compound.

Leger, J. M., Goursolle, M., Gadret, M. and Carpy, A. (**1978**). "Structure cristalline du sulfate de salbutamol [tert-butylamino-2 (hydroxy-4 hydroxyméthyl-3 phényl)-1 éthanol,0.5H2SO4]," Acta Cryst. B **34**, 1203–1208. CSD Refcode SALBUT.

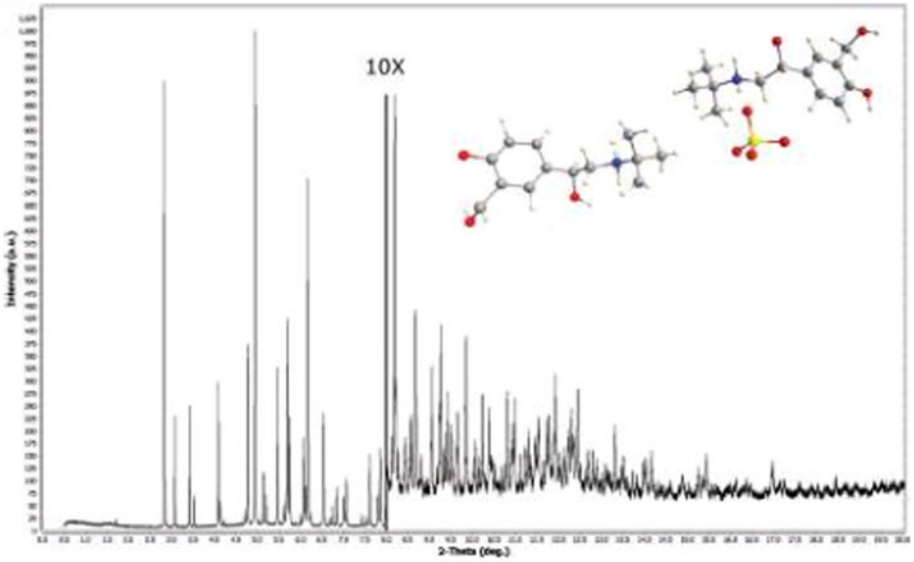

Figure 1. (Color online) Powder X-ray diffraction pattern of albuterol sulfate.

a)Author to whom correspondence should be addressed. Electronic mail: zhong@icdd.com
Data was submitted via Genie (http://www.icdd.com/websubmission/launch.html), the ICDD® Web Submission Page.

Figure 2. (Color online) The final result of a rapid communication submitted through the Genie program.

https://doi.org/10.1017/S0885715614000864 Published online by Cambridge University Press

that Genie can be run on all common operating systems, including Windows, the Mac OS X, and Linux. A Java Web Start application such as Genie also ensures that the correct JVM version is available and provides seamless version updating, thus guaranteeing compatibility. Also, each use of Genie will obtain the latest release from the ICDD. As a result, because of the use of Java Web Start, users do not need to concern themselves with installation or updating of software. They click on a URL in a web browser and all occurs seamlessly. Perhaps most importantly, Java Web Start applications run *outside* of the web browser, which avoids browser incompatibility issues. In short, Genie can be run on any reasonably modern computer with internet access.

Genie is flexible. Since it is not uncommon for powder data, crystal structures, reflection lists, etc. to be in separate files, Genie was written to allow users to quickly and easily gather their data from a *variety* of files types. This enables the merging of these data files to create a single unified submission for upload. Table I lists the currently supported file types.

B. The genie interface

Figure 1 shows the Genie interface. When using Genie, the user selects the intent of their submission: to submit to the ICDD PDF or to submit to the *Powder Diffraction Journal* as a rapid communication, as well [Figure 1(a)]. The information that is required for each submission depends on this choice. To guide the user, each of the tabs corresponding to required information for the selected publication type is colored red until the necessary data are entered.

Here we present an example in which information to be communicated are distributed in three files: the coordinates are in a CIF; the powder data, a GSAS file; and the reflection list, a columnar text file. In Figure 1(a), we see the user has started Genie and selected to have the submission considered for a rapid communication in the *Powder Diffraction Journal*, but no other input has yet been supplied. As a result, the Details, Powder Data, DI(*hkl*) List, and Publication Info tabs are colored red indicating that the data fields in these panes are required. Since neither a reference nor coordinates are required, the tabs for these panes remain brown. In Figure 1(b), we see that several fields in the *Details* pane where the user has entered, and as a result, the tab is now colored green indicating that nothing more is *required*. Since the DI(*hkl*) List, Powder Data, and Publication Info tabs are still red, the user knows that more information is required prior to submission.

Using Genie to read data in from different sources is easy. The menu bar clearly presents the user with options for reading in the powder data, reflection list and coordinates. Going to the *Read in Powder Data* menu and choosing GSAS, the user may read in the powder data from a GSAS format file. When the data are successfully read in, the *Powder Data* tab will be colored green.

The same steps are performed for reading in the reflection list [DI(*hkl*) List]. Since the reflection list is in a text file in this example, the user selects TXT under the *Read in DIList* menu. When the appropriate file is selected, the Genie text importer appears [Figure 1(c)]. This importer is a flexible tool for adapting too many different text file formats. As long as the data are in columns, the text importer can be used. Examining the data in the file as displayed in the File Preview pane, the user can see which values correspond to which columns. In this example, we see that the indices h, k, and l are in the first three columns, with 2θ and intensity following in the fourth and fifth columns. By entering the appropriate column numbers and pressing the "≫" GUI button, the data are properly extracted. At this point the user can examine what has been imported to ensure that items have been read in correctly.

Coordinates can also be imported in the same fashion by going to the *Read in Coordinates* menu and in this example selecting CIF. All the coordinates and anisotropic temperature factors, if appropriate, are then immediately imported from the selected CIF.

Figure 1(d) displays the *Publication Info* pane, which contains any author information, relevant keywords, and an abstract. Here all data have been filled in. Since all required information has been supplied, the Upload GUI button becomes active as the user can submit the collection of data to the ICDD when ready. Pressing that upload button causes the transfer to an ICDD server, which generates an automated response indicating the success of the upload.

The submission is then submitted for review by the ICDD scientific editors. The user will be contacted by the editorial staff and provided with a tracking number for their submission. The goal of the ICDD is to create a prompt and open dialog with the user, as the scientific editors may have questions concerning the submission. This ensures that the high quality of submissions to the PDF and *Powder Diffraction Journal* is maintained while also allowing the time from submission to publication to be minimal.

Figure 2 shows what the author can expect when a submission of powder data and coordinates has been accepted and published in the *Powder Diffraction Journal*. In this instance, the user received a publication in less than 3 months.

II. CONCLUSION

Genie, a new program for submitting data for publication to the PDF and as a rapid communication in the *Powder Diffraction Journal*, has been presented. Genie is a lightweight Java program that can read in and combine data from a variety of sources and subsequently upload the data to the ICDD for immediate review. Since Genie is written in Java, it is platform independent allowing the user to submit data from any computer including but not limited to Mac, Windows, and Linux. Genie can be readily accessed from the ICDD website at URL http://www.icdd.com/websubmission/launch.html

ACKNOWLEDGEMENT

Special thanks are due to Tom Blanton and Justin Blanton from the International Centre for Diffraction Data and Brian Toby of Argonne National Laboratory for their critical reading of this manuscript.

Bish, D. L. and Howard, S. A. (**1988**). "Quantitative phase analysis using the rietveld method," J. Appl. Crystallogr. **21**, 86–91.

Hanawalt, J. D. (**1983**). *History of the Powder Diffraction File, in Crystallography in North America – Apparatus and Methods*, ch. 2, pp 215–219. American Crystallographic Association, New York, NY.

ICDD (**2013**). PDF4+ 2013 (Database), edited by Dr. Soorya Kabekkodu, International Centre for Diffraction Data, Newtown Square, PA, USA.

ICDD (**2014**). PDF4+ 2014 (Database), edited by Dr. Soorya Kabekkodu, International Centre for Diffraction Data, Newtown Square, PA, USA.

Rietveld, H. M. (**1969**). "A profile refinement for nuclear and magnetic structures," J. Appl. Crystallogr. **2**, 65–71. doi: 10.1107/S0021889869006558