

International Tables for Crystallography

Volume G: Definition and exchange of crystallographic data

Sydney Hall and Brian McMahon, Editors

Published for the IUCr by Springer, 2005

ISBN 1-4020-3138-6, 594 + xii pages

This volume of *International Tables* must be aimed primarily at crystallographic programmers and people looking after the



maintenance and distribution of crystallographic data. The sections on the Molecular Information File (MIF) may be of interest to practical crystallographers wishing to save their noncrystallographic information in a digital format closely related to the ubiquitous CIF format, and the final part lists tools to help in the preparation of files for deposition or publication.

Much of the information it contains is available online from the IUCr website. The decision to publish a printed version reaffirms that while digital versions of texts are convenient for keyword searching and occasional reference, current hardware is nowhere near fast enough and current monitors are nowhere near big enough to permit effective browsing and research studies.

Part 1, a Historical Introduction, is more than just a history; it also provides an explanation for why the current crystallographic information files have a syntax that at first sight seems overcomplicated. It shows

the care and insight that went into designing and then revising the data dictionaries, which, in effect, are an ordered mechanism for representing almost all

crystallographic data. The final section, the relationship between CIF and XML, shows that (fortunately) these two concepts have sufficient in common to make data transfer from one system to the other relatively simple.

Part 2, Concepts and Specifications, develops the ideas introduced in Part 1, and is probably the key section for programmers to study. Section 2.1.3.1, which describes the formal definitions of <blank> (ASCII 32, 11 and 9), <terminate> (ASCII 10, 12 and 13) and <wspace> and their association with the mark-up characters ' , " and ; to define text strings, will be important to programmers, as will the portability and archival issues described in Section 2.2.4. For example, to be fully forward compliant, programmers writing codes to read CIF files will need to process text lines consisting of up to 2048 characters (as opposed to the original definition of 80 characters). The specification of the Crystallographic Binary File (CBF/imgCIF) describes how figures and drawings, together with diffraction images, can be included into CIF-like files. Section 2.5 (core CIF dictionary definition language, DDL1) and Section 2.6 (relational dictionary definition language, DDL2) explain the concepts behind the many data items described later in the volume.

Part 3, CIF Data Definition and Classification, is 130 pages of carefully written explanations of the practical issues in defining data items, covering the core data, powder diffraction, modulated and composite structures, macromolecular data, image data and symmetry data. The CIF concept enables individuals or laboratories to define and register data items for their own use, and include them in data files. These will be

read but ignored by a properly compliant CIF-reading program. However, there are very many data items that will be in very widespread use, and which should thus have carefully thought out published definitions. The IUCr Committee for the Maintenance of the CIF Standard (COMCIFS) has the task of producing and maintaining these fundamental definitions.

Part 4, Data Dictionaries, occupies almost half the volume, and gives the current definitions of all data items to be found in the various data dictionaries. This is the section that most benefits from having an electronic version available online, as that can be quickly searched without one having to know the exact data name.

Finally, Part 5, Applications, provides advice for programmers and CIF users. It details libraries of useful CIF facilities that can be built into other applications, and also complete applications that can be downloaded or accessed, mostly free of charge, for editing, validating and manipulating CIF. It is the section that will most rapidly become out of date. The volume also includes a CD containing most of the tabular material, software libraries and end-user applications.

The CIF format has more or less become the international standard for the deposition and publication of crystallographic results, and increasingly application programs will accept CIF as an input format. Most practising structure analysts will expect their normal software systems to create the bulk of a CIF for each structure. However, most CIFs will require some additional manual editing — increasingly so as journals become willing to accept whole papers in CIF format.

International Tables Volume G will be an invaluable reference to help ensure that edits are compliant, and may also encourage crystallographers to store other information along with

the minimal crystallographic details. Every active crystallography group should have a copy of this book.

“Every active crystallography group should have a copy of this book”

Reviewed by David Watkin

Chemical Crystallography Laboratory, U. of Oxford, UK

For more information about *International Tables*, including how to order a print copy of Volume G, please visit <http://it.iucr.org>.



International Tables are now available online as a full set of volumes through Springer; please visit <http://www.springer.com/978-1-4020-5259-0>.