# 5.7. Small-molecule crystal structure publication using CIF

By P. R. Strickland, M. A. Hoyland and B. McMahon

## 5.7.1. Introduction

The International Union of Crystallography (IUCr) has always understood the importance of the accurate reporting of numerical results, and as far back as its early sponsorship of the Standard Crystallographic File Structure (Brown, 1983, 1988) the IUCr has explored the use of exchange files in publishing (see Chapter 1.1). In 1991, when the first draft of the CIF standard was nearing completion, the main journal of the IUCr for reporting crystal structures, *Acta Crystallographica Section C: Crystal Structure Communications* (hereafter *Acta Cryst. C*), consisted of a collection of concise reports of crystal and molecular structures presented in a standard format that would lend itself well to computerized markup and typesetting from an appropriate input file format. It seemed natural, therefore, to use this journal to test the new draft CIF standard and to develop techniques for machine-based checking of structural data along with the new methods for submitting, typesetting and distributing a crystal-structure report in electronic format. Although adopting a novel data-exchange format for the submission and handling of research papers might have seemed a radical and audacious development, the potential benefits in terms of accuracy and speed of publication were clear.

In parallel with the publication of the CIF standard (Hall *et al.*, 1991), an *Editorial* and revised *Notes for Authors* in *Acta Cryst. C* described the new route to publication using CIFs and invited the crystallographic community to cooperate in this innovative practice. The same issue of the journal contained the first paper to be published by this route (Willis *et al.*, 1991).

This first paper was the outcome of a testing phase which involved considerable interaction with the authors. The first unsolicited article to be submitted in CIF format appeared in the February 1992 issue of *Acta Cryst. C*. A few more were submitted during 1992, the number gradually increasing through the following year. Authors quickly adapted to the compartmentalized style of text entries and by the beginning of 1994 the level of CIF submissions allowed the journal to introduce a production stream that promised faster publication times for articles submitted electronically as CIFs. By the beginning of 1996, it became journal policy to accept *only* electronic submissions in CIF format.

The IUCr was not the only publisher to introduce the submission of structure reports in machine-readable form. In 1990, *Zeitschrift für Kristallographie*, published by R. Oldenbourg Verlag, introduced a new section for the publication of short inorganic and small-molecule structural papers with minimal commentary. To submit a report to this section, the author would use the output file from the refinement program *SHELX76* (Sheldrick, 1976) (at that time a *de facto* exchange standard on account of its widespread distribution), which was processed by a specially developed program *CASTOR* to create a self-contained file for use in publication. When CIF was introduced, it was also accepted as a submission format for this section of *Zeitschrift*. The section flourished and in 1997 it became a separate journal, *Zeitschrift*

*für Kristallographie – New Crystal Structures*. CIF is now the standard submission format for this journal as well as for *Acta Cryst. C*.

In an era dominated by information retrieval *via* the world wide web, it is easy to forget that these innovations in crystallographic publishing predated the http protocol and the universal availability of graphical browsers. However, the independently developed but well defined CIF exchange standard proved easy to integrate with the publication procedures developed for electronic journals. The current delivery formats available to journals like *Acta Crystallographica* and *Zeitschrift für Kristallographie* are HTML and PDF. Nevertheless, the original CIF data are still accessible, and allow readers to visualize structures interactively in three dimensions or perform their own analyses of structural models.

The highly automated submission, checking and publication procedures of *Acta Cryst. C* and the online-only journal *Acta Crystallographica Section E: Structure Reports Online* (hereafter *Acta Cryst. E*) are described in detail in Section 5.7.2 as a case study for the publication of structure reports that are highly ordered in format. However, there are only a few journals that report detailed crystal structures and they represent a very specialized field of publishing. Section 5.7.3 discusses publications in which the reporting of structural data is only a minor or supplementary element of the article. It will become apparent that many of the considerations behind the design of a workflow for handling data-rich papers are also relevant to maximizing the value of data presented in or referenced by any scientific publication.

## 5.7.2. Case study: the fully automated reporting of small-unit-cell crystal structures

This section describes the route to publication of a small-molecule or inorganic single-crystal structure in *Acta Cryst. C* or *E* from the perspective of an author.

### 5.7.2.1. Assembling the complete article

For many authors the generation of a CIF suitable for publication is quite straightforward, since diffractometer software and structure solution and refinement packages have all been capable of writing or reading the CIF format for some time. In some highly integrated systems, the entire experimental, analysis and report-generating pathway may be controlled through a common user interface.

In other cases, different components must be collected from different sources and merged together, either by software utilities or, in the worst case, by hand-editing. It is a useful feature of the text-based CIF format that it can be modified by text editors or in certain word-processing modes; indeed, this was the only way in which the earliest CIF-based papers could be constructed. However, significant expertise and understanding of the technical details of the file format are needed to produce hand-edited files that are totally free from error. Authors are now encouraged to use software designed to help them create complete and error-free files (*e.g.* the *enCIFer* and *CIFEDIT* editors described in Chapter 5.3).

A complete structure communication comprises the following components.

**references**