Acta Cryst. (2002). A**58** (Supplement), C258

CIF AND CHEMICAL INFORMATION SYSTEMS

B. McMahon
International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, UK

The Crystallographic Information File (CIF) has for over a decade been the standard mechanism for electronic information interchange in crystallography. Its success is due partly to the specialized nature of the information exchanged by the coherent community of crystallographers. By contrast, chemistry in general embraces a much larger community with many different requirements; several initiatives in chemical information representation and exchange have been unsuccessful. Increasingly, however, the chemical community sees eXtensible Markup Language (XML) as a suitable vehicle for information exchange. A Task Group of the International Union of Pure and Applied Chemistry (IUPAC) is currently charged with making effective use of XML. It proposes to acquire from the body supervising XML development the authority to define markup tags with chemical content, and to identify specific topic areas in which suitable tags are to be defined. Chemical Markup Language (CML) is an example of a chemical XML document type definition. The format of XML is different from the CIF syntax, but mechanistic transformation between file formats is becoming more common as abstract data models are increasingly refined. The creation of suitably meaningful tags will however be a formidable challenge. Defining the tags in a new data dictionary for a CIF extension can occupy a group of experts for several years and result in a few dozen new identifiers. Existing CIF and related molecular information file (MIF) dictionaries provide a source for identifiers that, if adopted in a general chemical context, could form a bridge between crystallographic and chemical information exchange standards.

Keywords: CIF XML CHEMICAL INFORMATION