International Tables for Crystallography (2006). Vol. G, Section 3.4.2, pp. 131–132.

3.4. Classification and use of modulated and composite structures data

BY G. MADARIAGA

3.4.1. Introduction

Aperiodic structures do not have lattice periodicity, but do show long-range order. Their diffraction patterns exhibit sharp Bragg peaks that have to be indexed using more than three Miller indices. All aperiodic structures can be classified into one of three groups: incommensurately modulated structures, incommensurate composite structures and quasicrystals. It is the purpose of the modulated and composite structures CIF dictionary (msCIF dictionary), presented in Chapter 4.3, to provide machine-readable descriptions of the first two of these groups.

Modulated structures represent the simplest case and are described by periodic distortions of an underlying (reference) crystal structure. The distortions involve one or several atomic parameters – positions, occupation probability or thermal parameters (the term *displacement parameters* is ambiguous in this context) – and their periodicity may be commensurate or incommensurate with the lattice of the reference structure. The first case corresponds to a three-dimensionally periodic crystalline phase, whereas the second case defines an incommensurate structure. In both cases, the diffracted intensities can be divided into two groups: the prominent main reflections, which are located in a three-dimensional reciprocal lattice (that of the reference structure), and additional (generally) weaker satellite reflections situated at points determined by the wave vectors of the modulations. Strictly speaking, the number of parameters needed to describe the atomic modulations of an incommensurate structure is infinite. In practice, however, the number of structural parameters is often drastically reduced owing to the strong hierarchy (demonstrated by the discreteness of the diffraction diagrams) that exists among them.

Composite structures can be thought of as being built of two or more periodic subsystems whose lattices are mutually incommensurate. Therefore, the set of main reflections no longer defines a unique three-dimensional reciprocal lattice. Moreover, the interaction between the different subsystems provokes modulations and, as a consequence, the presence of satellite reflections in the diffraction diagram.

Quasicrystals not only lack three-dimensional lattice periodicity but also show noncrystallographic symmetry.

The methodology for solving aperiodic crystal structures has been well developed since the introduction of the concept of superspace (de Wolff, 1974; de Wolff *et al.*, 1981). Superspace allows the recovery of the periodicity and a simple description of the symmetry of quasicrystalline structures in a higher-dimensional space. The real aperiodic structure is recovered from the superspace through appropriate three-dimensional sections. The dimension of the superspace $(3 + d)$ is equal to the number of reciprocal vectors needed to index the whole diffraction pattern of the quasicrystalline structure. More information on the superspace approach can be found, for example, in van Smaalen (1995) and Janssen *et al.* (2004).

The success of this concept relies above all on the systematic description of the symmetry of the aperiodic materials using superspace groups. Superspace groups for the simplest (but most common) case of one-dimensional modulated structures are tabulated in Janssen *et al.* (2004).

Within the superspace approach, all the aperiodic atom positions are embedded in dense sets (atomic domains) in the $(3 + d)$-dimensional unit cell of the associated periodic structure in superspace. They are parallel on average to the 'internal' (or 'perpendicular') space, which is a d-dimensional subspace chosen to be orthogonal to the real (physical or parallel) space. The three-dimensional structure is then a section of the $(3+d)$-dimensional structure parallel to the real space. The atomic domains are distorted along the internal space by the modulation functions. In many incommensurate structures and composites, the atomic domains are continuous periodic functions along the internal subspace and are parameterized by Fourier series. Some compounds, however, need to be described using discrete atomic domains whose parameterization is more complicated and will be discussed in Section 3.4.2. Note that commensurate structures can also be included within the superspace approach. The difference between incommensurate and commensurate modulations is that for commensurate modulations only a finite number of values of the atomic modulation functions are relevant.

The number of modulated and composite structures solved with software that uses the superspace formalism has grown rapidly and in many cases the determination of the structures of such systems is now almost routine. A standard for the description of incommensurate modulated structures has been established by the Commission on Aperiodic Crystals of the International Union of Crystallography (Chapuis *et al.*, 1997) using a checklist that is easily extensible to composite crystals. However, there was until recently no standard way to represent these structures electronically. Structural databases tend to contain only a brief reference to the modulated character of the structure and it has not been possible to transmit, archive and retrieve information about modulated and composite structures as efficiently as for normal crystal structures. Extending the core CIF dictionary to form a modulated and composite structures CIF (msCIF) dictionary seemed the appropriate way to deal with these problems, and has the additional benefits derived from the use of a well tested standard for which several tools have been developed.

In the case of quasicrystals, however, although the theoretical foundations seem to be well established, the determination of accurate models requires a combination of different strategies and techniques. The parameterization of the atomic domains with physical meaning is far from being an automated procedure (Cervellino *et al.*, 2002) and some of the existing models are now disputed. Major problems arise from sample quality, intrinsic disorder and rather low data-to-parameter ratios (Haibach *et al.*, 2000). Quasicrystals are not covered by the msCIF dictionary detailed in Chapter 4.3.

3.4.2. Dictionary design considerations

The CIF dictionary for modulated and composite structures (msCIF dictionary) is an extension of the core CIF dictionary

Affiliation: GOTZON MADARIAGA, Departamento de Física de la Materia Condensada, Facultad de Ciencia y Tecnología, Universidad del País Vasco, Apartado 644, 48080 Bilbao, Spain.

Table 3.4.2.1. *Category groups defined in the msCIF dictionary*

The groups are listed in the order in which they are described in this chapter.

Section	Category group	Subject covered
(*a*) Experimental measurements		
3.4.3.1.1	CELL	The unit cell, especially cell subsystems for composite structures
3.4.3.1.2	DIFFRN	Multi-dimensional diffraction pattern
3.4.3.1.3	EXPTL	Details of the experiment specific to modulated and composite structures
(*b*) Analysis		
3.4.3.2.1	REFINE	Refinement procedures
3.4.3.2.2	REFLN	Reflection measurements indexed in higher-dimensional space
(*c*) Atomicity, chemistry and structure		
3.4.3.3.1 to 3.4.3.3.3	ATOM	Atom sites in a modulated structure
3.4.3.3.4	GEOM	Geometry of a structure in superspace
3.4.3.3.5	SYMMETRY	Symmetry information
(*d*) File metadata		
3.4.3.4	AUDIT	The structure of the CIF

(Chapter 4.1). This means that the dictionary defines items that are basically related to single-crystal data. The close relationship between the msCIF and core dictionaries has led to synergies between and benefits for both dictionaries.

The design of the msCIF dictionary had two objectives: (i) it should be as functional as possible, *i.e.* as little information in an msCIF as possible should be given as unstructured text; (ii) it should be possible to include even the oldest modulated and composite structures in an msCIF, even if the way they were reported did not follow the guidelines used now.

There were two major difficulties in implementing the msCIF dictionary. Firstly, the number of additional wave vectors used to index a diffraction diagram is theoretically not limited. Secondly, a CIF containing information about a modulated or composite structure should, in general, be composed of several (related) data blocks. As CIF definitions do not at present include vectors or matrices as distinct types of data values, an arbitrary upper limit of 11 was assigned to the dimension of superspace to limit the number of new data names. Linking between data blocks is handled by using recommended values for items in the AUDIT and AUDIT_LINK categories, like those used in the powder CIF (pdCIF) dictionary (Chapters 3.3 and 4.2).

An additional problem arises when special (ideal) modulation functions are considered. Although periodic modulations are normally parameterized by Fourier series, in certain cases it is convenient to use discontinuous functions which lead to a severe reduction in the number of structural parameters. The shape of these functions is not restricted and new materials could require new functions. Given that it is not possible at this moment to define logical or mathematical relations between data values within a CIF [although an initiative for including algorithms in the definitions of CIF dictionaries has been proposed by Spadaccini *et al.* (2000)], general functions cannot be defined and therefore the type of special functions included in the msCIF dictionary are those implemented in the most widely used program, *JANA2000* (Petříček & Dušek, 2000). They only apply to one-dimensional modulations and are sawtooth displacive functions and occupational crenel functions. Both functions define discontinuous occupational atomic domains and are normally combined with (smoother) atomic modulation functions (involving atom positions and/or thermal parameters) that are expressed by Fourier series. Because of the discreteness of the atomic domains, members of the set of harmonic functions used to expand these series are no longer mutually orthogonal as they are only defined within each atomic domain and not in the (internal space) interval $[0, 1]$. As a consequence, severe correlation effects among the coefficients of the Fourier series are expected. A solution for this problem lies in the selection and orthogonalization of a set of basic functions (Petříček *et al.*, 1995). The atomic modulation functions are then expressed as linear combinations of an orthogonal basis whose elements are specific combinations of harmonic functions. Discontinuous atomic domains are being increasingly used in composite materials, in some cases revealing that considering such materials as composites or modulated structures is a matter of convenience (Elcoro *et al.*, 2003; Pérez-Mato *et al.*, 2003).

CIFs that conform to the msCIF dictionary are highly itemized for the human reader, but have a strong relational structure even though the dictionary itself is written in DDL1.

The major drawbacks of the dictionary are:

(i) Items describing the superspace symmetry should be reconsidered and perhaps included within the symmetry CIF (symCIF) dictionary (Chapters 3.8 and 4.7).

(ii) There is still some information in an msCIF that cannot be interpreted by a computer [*e.g.* rigid rotations and translations around (along) noncrystallographic axes cannot be parsed, since the description of the axes is textual].

(iii) A full description of the modulation in terms of orthogonalized functions (used when the atomic domains are discrete) is not supported yet.

3.4.3. Arrangement of the dictionary

The msCIF dictionary detailed in Chapter 4.3 includes 19 new categories. Another 18 already exist in the core CIF dictionary, but include new items (16) or revised definitions (2). The category structure of the msCIF dictionary is summarized in Table 3.4.2.1 and is listed in full in Appendix 3.4.1. The appendix also lists for each category the section of this chapter in which the category is described.

Many of the modifications to categories that already exist in the core CIF dictionary result from the need to use more than three integer indices to label the diffracted intensities (in the cases of CELL, DIFFRN_REFLN, DIFFRN_REFLNS, DIFFRN_STANDARD_REFLN and EXPTL_CRYSTAL_FACE) or the need to use superspace symmetry (in the cases of GEOM_ANGLE, GEOM_BOND, GEOM_CONTACT, GEOM_TORSION, SPACE_GROUP and SPACE_GROUP_SYMOP). Apart from the categories that describe the atomic modulation functions, there are two that are specific to composite structures (CELL_SUBSYSTEM and CELL_SUBSYSTEMS).

The rest of this section summarizes the contents of the dictionary, organized by categories within the functional groups outlined in Table 3.1.10.1. As in the other chapters in this part of the volume, the classification is under the headings *Experimental measurements* (Section 3.4.3.1), *Analysis* (Section 3.4.3.2), *Atomicity, chemistry and structure* (Section 3.4.3.3) and *File metadata* (Section 3.4.3.4). The msCIF dictionary adds no new data items concerned with the publication or reporting of structures to those already present in the core CIF dictionary.

The data items within each category are listed in the detailed commentary below. Where relevant, data items that represent a unique identifier for a looped list ('category keys') are listed first and are marked by a bullet (•). The remaining data items in each category are listed alphabetically.