

## 3.4. CLASSIFICATION AND USE OF MODULATED AND COMPOSITE STRUCTURES DATA

Data items revised in these categories are as follows:

- (a) AUDIT  
     \_audit\_block\_code  
 (b) AUDIT\_LINK  
     \_audit\_link\_block\_code

The core dictionary definitions of these items are revised in order to formalize the relationships between multiple data blocks representing reference and modulated structures. Guidance is provided in the msCIF dictionary on how to label data blocks in a way that makes their mutual relationships clear.

### 3.4.4. Use of the msCIF dictionary

In this section, some of the capabilities of the dictionary will be demonstrated using simple examples. More detailed examples can be found at <http://www.iucr.org/iucr-top/cif/ms> and on the CD-ROM accompanying this volume.

#### 3.4.4.1. Description of reciprocal space

Modulated and composite structures need more than three reciprocal vectors in order to index the whole set of reflections with integer numbers. Hence a diffraction vector is written as

$$\mathbf{H} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^* + m_1\mathbf{q}_1 + \dots + m_d\mathbf{q}_d, \quad (3.4.4.1)$$

where the notation has been chosen according to the core CIF dictionary. In the case of a modulated structure,  $\mathbf{a}^*$ ,  $\mathbf{b}^*$  and  $\mathbf{c}^*$  are the reciprocal vectors of the reference structure (and therefore  $h$ ,  $k$  and  $l$  index the main reflections).  $\mathbf{q}_1, \dots, \mathbf{q}_d$  are the modulation wave vectors. They are three-dimensional vectors with some irrational component (if the modulated structure is incommensurate) in the lattice spanned by  $\mathbf{a}^*$ ,  $\mathbf{b}^*$  and  $\mathbf{c}^*$ .  $d$  is the dimension of the modulation. In the case of composite structures, the diffraction pattern can be indexed using  $3 + d$  (arbitrarily selected) vectors  $\mathbf{a}_k^*$  ( $k = 1, \dots, 3 + d$ ).  $\mathbf{a}_1^*$  ( $\equiv \mathbf{a}^*$ ),  $\mathbf{a}_2^*$  ( $\equiv \mathbf{b}^*$ ) and  $\mathbf{a}_3^*$  ( $\equiv \mathbf{c}^*$ ) normally span the reciprocal lattice of the main reflections of one of the substructures (notice that this is only one particular, but highly intuitive, choice). The remaining  $d$  vectors with  $k = 4, \dots, d$  are the wave vectors of the modulation [ $\mathbf{q}_1, \dots, \mathbf{q}_d$  in equation (3.4.4.1)].

In a composite structure, the  $(3 + d)$ -dimensional reciprocal basis of the subsystem  $\nu$  is determined by a  $(3 + d) \times (3 + d)$  matrix  $W^\nu$  [see van Smaalen (1995) and references therein]:

$$\mathbf{a}_i^{*\nu} = \sum_{k=1}^{3+d} W_{ik}^\nu \mathbf{a}_k^*, \quad i = 1, \dots, 3 + d, \quad (3.4.4.2)$$

where the subscripts  $i = 1, 2$  and  $3$  label the reciprocal vectors  $\mathbf{a}^{*\nu}$ ,  $\mathbf{b}^{*\nu}$  and  $\mathbf{c}^{*\nu}$ , and  $i = 4, \dots, d$  label the wave vectors of the modulation expressed as linear combinations of  $\mathbf{a}^{*\nu}$ ,  $\mathbf{b}^{*\nu}$  and  $\mathbf{c}^{*\nu}$ .

The simplest case corresponds to a one-dimensional ( $d = 1$ ) modulated structure. Consider for example the incommensurate phase of  $\text{K}_2\text{SeO}_4$ . The wave vector of the modulation can be chosen to be  $\mathbf{q}_1 = \alpha\mathbf{a}^*$ . Relevant information about the diffraction pattern of this compound is expressed using both the core CIF and msCIF dictionaries as shown in Example 3.4.4.1.

A more complicated example is the composite structure  $(\text{LaS})_{1.14}\text{NbS}_2$ . The two mutually incommensurate subsystems (along the  $a$  axis) are (van Smaalen, 1991)  $\text{NbS}_2$  ( $\nu = 1$ ) and  $\text{LaS}$  ( $\nu = 2$ ). The reciprocal basis can be chosen to be  $\mathbf{a}_1^* = \mathbf{a}^{*1}$ ,  $\mathbf{a}_2^* = \mathbf{b}^{*1}$ ,  $\mathbf{a}_3^* = \mathbf{c}^{*1}$  and  $\mathbf{a}_4^* = \mathbf{a}^{*2}$ . For this particular choice, the two  $W$  matrices [see equation (3.4.4.2)] are

Example 3.4.4.1. msCIF description of the diffraction pattern of a one-dimensional modulated structure.

```

_exptl_crystal_type_of_structure      'mod'
_cell_reciprocal_basis_description
; a*,b*,c* (reciprocal basis spanning the lattice of
main reflections), q modulation wave vector.
;

_diffn_symmetry_description
; The whole diffraction pattern shows orthorhombic
symmetry. The following extinction rules were
detected:
    0k10  k+1=odd
    h0lm  h+m=odd
    hk0m  m=odd
    h00m  h,m=odd
Superspace group: P:Pnam:-1ss
;

_diffn_reflns_satellite_order_max      1

_diffn_reflns_theta_max                 40.14
_diffn_reflns_theta_min                 3.32
_diffn_reflns_limit_h_max               8
_diffn_reflns_limit_k_max              18
_diffn_reflns_limit_l_max              10
_diffn_reflns_limit_index_m_1_max      1
_diffn_reflns_limit_h_min               0
_diffn_reflns_limit_k_min               0
_diffn_reflns_limit_l_min               0
_diffn_reflns_limit_index_m_1_min      -1

# Modulation wave vector
loop_
_cell_wave_vector_seq_id
_cell_wave_vector_x
    1          0.318(5)
    
```

$$W^1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad W^2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

This information is transcribed to CIF format as shown in Example 3.4.4.2. (Note that the default values for the wave vector components and the elements of  $W$  are 0.)

#### 3.4.4.2. Description of symmetry

The symmetry of a modulated or composite structure is described by a superspace group which leaves the  $(3 + d)$ -dimensional embedding of the structure invariant. Superspace is built of two orthogonal subspaces and both of them are kept invariant separately by the superspace symmetry operations. (In reciprocal space this means that, for these structures, main reflections and satellite reflections are never transformed into one another by superspace symmetry operations.) Consequently, superspace groups are not general  $(3 + d)$ -dimensional space groups. The standard notation for superspace groups only covers the one-dimensional superspace groups, which are listed in Janssen *et al.* (2004). As a consequence, msCIFs must include a list of all the symmetry operations in an  $(x, y, z)$  format (using as symbols  $x_1 \dots x_{3+d}$ ) similar to that used in the core CIF dictionary. Superspace-group names for one-dimensional structures can be expressed either according to Janssen *et al.* (2004) or according to the original notation of de Wolff *et al.* (1981). Alternative names or higher-dimensional superspace groups can also be included, but not parsed.

### 3. CIF DATA DEFINITION AND CLASSIFICATION

Example 3.4.4.2. Representation of two mutually incommensurate subsystems.

```
_exptl_crystal_type_of_structure      'comp'
_cell_subsystems_number              2
_cell_modulation_dimension           1

_cell_reciprocal_basis_description
; a1*,b1*,c1* (reciprocal basis spanning the
reciprocal basis of the first subsystem),
a2* reciprocal axis corresponding to the
second subsystem.
;

loop_
_cell_subsystem_code
_cell_subsystem_description
_cell_subsystem_matrix_W_1_1
_cell_subsystem_matrix_W_1_4
_cell_subsystem_matrix_W_2_2
_cell_subsystem_matrix_W_3_3
_cell_subsystem_matrix_W_4_1
_cell_subsystem_matrix_W_4_4
NbS2 '1st subsystem' 1 0 1 1 0 1
LaS  '2nd subsystem' 0 1 1 1 1 0

_diffn_symmetry_description
; The whole diffraction pattern shows orthorhombic
symmetry. The following extinction rules were
detected:
hk0m h+k=odd
hk0m m=odd
Extinction rules are compatible with the superspace
groups:
P:Fmmm:-11s (\a,0,0)
P:Fm2m:-1-1s (\a,0,0)
;
```

Example 3.4.4.3. Symmetry description of a superspace group.

```
_space_group_ssg_name_WJJ          'P:P n a m:-1 s s'
_space_group_ssg_name_IT           'P n m a (\a 0 0) 0 s s'

loop_
_space_group_symop_ssg_id
_space_group_symop_ssg_operation_algebraic
1 x1,x2,x3,x4
2 1/2+x1,1/2-x2,1/2-x3,x4
3 1/2-x1,1/2+x2,-x3,1/2-x4
4 -x1,-x2,1/2+x3,1/2-x4
5 -x1,-x2,-x3,-x4
6 1/2-x1,1/2+x2,1/2+x3,-x4
7 1/2+x1,1/2-x2,x3,1/2+x4
8 x1,x2,1/2-x3,1/2+x4
```

In the particular case of  $K_2SeO_4$ , the superspace group is  $P_{1s}^{Pnma}$  (de Wolff *et al.*, 1981) or  $Pnma(\alpha 00)0s$  (Janssen *et al.*, 2004). This information would appear in a CIF as shown in Example 3.4.4.3.

#### 3.4.4.3. Description of the structure

A modulated structure is described by a reference periodic structure and the atomic modulation functions. Such functions are periodic and are normally expanded as Fourier series. The modulated parameters may apply to the atom positions (displacive modulation), the site occupancies (occupational modulation) and/or the temperature factors. In composite structures, each substructure is referred to the crystallographic basis defined by the  $W$  matrices [see equation (3.4.4.2)]. The simplest case corresponds to

Example 3.4.4.4. Atomic displacements as translations and rigid rotations.

```
loop_
_atom_site_description
_atom_site_label
_atom_site_occupancy
_atom_site_fract_x
_atom_site_fract_y
_atom_site_fract_z
; SeO4 is a tetrahedral group formed by the atoms
Se, O1, O2, O3 and O1p. O1p is related to O1 by the
mirror plane perpendicular to the c axis. The
coordinates given by _atom_site_fract_ correspond
to the centre of mass.
;
SeO4 1.0 0.22346 0.41868 0.250000
. Se 1.0 0.22216(4) 0.41861(3) 0.250000
. K1 1.0 0.1670(1) 0.08069(7) 0.250000
. K2 1.0 -0.0013(1) 0.70766(4) 0.250000
. O1 1.0 0.2890(3) 0.3413(2) 0.0277(3)
. O2 1.0 0.3130(4) 0.5621(2) 0.250000
. O3 1.0 0.0092(4) 0.4304(3) 0.250000

_refine_ls_mod_func_description
; Displacive modulation. Fourier series. Modulation
of SeO4 group described in terms of rigid
translations and rotations.
;

loop_
_atom_site_Fourier_wave_vector_seq_id
_atom_site_Fourier_wave_vector_x
_atom_site_Fourier_wave_vector_description
1 0.318(5) 'First harmonic'

loop_
_atom_site_displace_Fourier_id
_atom_site_displace_Fourier_atom_site_label
_atom_site_displace_Fourier_axis
_atom_site_displace_Fourier_wave_vector_seq_id
K1_z_1 K1 z 1
K2_z_1 K2 z 1
SeO_z_1 SeO4 z 1

loop_
_atom_site_displace_Fourier_param_id
_atom_site_displace_Fourier_param_cos
_atom_site_displace_Fourier_param_sin
K1_z_1 0.0080(4) -0.0106(5)
K2_z_1 0.0159(4) 0.0071(6)
SeO_z_1 -0.0089(2) -0.0058(2)

loop_
_atom_site_rot_Fourier_id
_atom_site_rot_Fourier_atom_site_label
_atom_site_rot_Fourier_axis
_atom_site_rot_Fourier_wave_vector_seq_id
SeO_x_1 SeO4 x 1
SeO_y_1 SeO4 y 1

loop_
_atom_site_rot_Fourier_param_id
_atom_site_rot_Fourier_param_cos
_atom_site_rot_Fourier_param_sin
SeO_x_1 -4.2(1) 0.91(3)
SeO_y_1 4.3(1) 0.
```

a one-dimensional displacive modulated structure. In this case, the atomic modulation functions are given by

$$u_{\alpha}^{\mu} = \sum_{n=1}^{\infty} U_{n\alpha}^{\mu} \cos(2\pi n\mathbf{q} \cdot \mathbf{r} + \varphi_{n\alpha}^{\mu}), \quad (3.4.4.3)$$

where  $(U_{n\alpha}^{\mu}, \varphi_{n\alpha}^{\mu})$  is the complex amplitude of each Fourier term;  $\mu$  labels the atoms;  $\alpha = x, y, z$ ;  $\mathbf{r}$  is the average atom position; and  $n\mathbf{q}$  represents the successive harmonics of the modulation. At present, displacive modulations along axes other than  $a$ ,  $b$  and  $c$  can be calculated with the restriction stated in Section 3.4.2.

## Example 3.4.4.5. Relationship between data blocks in an msCIF.

```

_audit_block_code      1997-07-24|LaSNbS2|G.M.|
loop_
_audit_link_block_code
_audit_link_block_description
. 'common experimental and publication data'
1997-07-24|LaSNbS2|G.M.|_REFRNC
'reference structure (global data)'
1997-07-21|LaSNbS2|G.M.|_MOD
'modulated structure (global data)'
1997-07-24|LaSNbS2|G.M.|_REFRNC_NbS2
'reference structure (1st subsystem)'
1997-07-21|LaSNbS2|G.M.|_MOD_NbS2
'modulated structure (1st subsystem)'
1997-07-24|LaSNbS2|G.M.|_REFRNC_LaS
'reference structure (2nd subsystem)'
1997-07-21|LaSNbS2|G.M.|_MOD_LaS
'modulated structure (2nd subsystem)'

```

Atomic displacements can also be expressed as rigid rotations and translations. The incommensurate phase of  $K_2SeO_4$  is one-dimensional displacive and sinusoidal (at least over a wide range of temperature), *i.e.*  $n = 1$  in equation (3.4.4.3). The tetrahedral  $SeO_4$  groups behave as rigid bodies. Symmetry considerations restrict the possible translations to occur along  $c$  where the only allowed rigid rotations are around the  $a$  and  $b$  axes. The incommensurate structure is then expressed as shown in Example 3.4.4.4.

Alternatively, equation (3.4.4.3) can be expressed as a Fourier series with real amplitudes. This form is also covered by the msCIF dictionary. Note that there is a global phase, which is irrelevant in the incommensurate case but fixes the space group if the modulated structure is commensurate. Global phases are also defined in the dictionary.

#### 3.4.4.4. Block pointers

CIFs for modulated and (in particular) composite structures often need several linked data blocks. For example, the average structure of a modulated phase refined using only the main reflections could be reported in one data block and then the results of the refinement of the complete structure using all the reflections could be added as a separate data block at a later stage. Similarly, each substructure in a composite structure can be fully described in a separate data block. It is therefore essential to keep all these blocks together, since all of them together describe the composite structure. Block pointers are described in the core CIF dictionary, but they need to be used with particular care in msCIFs. Therefore, the msCIF dictionary gives additional guidelines for naming the data blocks, as the aim (as in pdCIF) is to define unique names. Example 3.4.4.5 shows how the structure of  $(LaS)_{1.14}NbS_2$  can be represented using several data blocks. In this example, the block named 1997-07-24|LaSNbS2|G.M.| is linked to those in the loop below.

#### 3.4.4.5. Other information

The msCIF and core CIF dictionaries also include items for  $R$  factors for different reflection classes (for example, for the main reflections and for the satellite reflections; see Section 3.2.3) and for describing the variation of interatomic distances due to the modulation. msCIF can be used to create a complete record of the determination of a modulated or composite structure suitable for submitting as an article to an academic journal or for archiving in a database. The obvious next step in the development of msCIF is to adapt or write software that can filter and manage msCIF data and that can generate msCIFs automatically.

### Appendix 3.4.1

#### Category structure of the msCIF dictionary

Table A3.4.1.1 shows all categories that are represented in the msCIF dictionary.

Table A3.4.1.1. Categories in the modulated structures CIF dictionary

Numbers in parentheses refer to the section of this chapter in which each category is described in detail.

ATOM group (§3.4.3.3)	AUDIT group (§3.4.3.4)
ATOM.SITE (§3.4.3.3.1(a)) ¶	AUDIT (§3.4.3.4(a)) ‡
ATOM.SITE.DISPLACE.FOURIER (§3.4.3.3.2(a))	AUDIT.LINK (§3.4.3.4(b)) ‡
ATOM.SITE.DISPLACE.FOURIER. PARAM (§3.4.3.3.2(b))	CELL group (§3.4.3.1.1)
ATOM.SITE.DISPLACE.SPECIAL.FUNC (§3.4.3.3.3(a))	CELL (§3.4.3.1.1(a)) ¶
ATOM.SITE.FOURIER.WAVE.VECTOR (§3.4.3.3.2(c))	CELL.SUBSYSTEM (§3.4.3.1.1(b))
ATOM.SITE.OCC.FOURIER (§3.4.3.3.2(d))	CELL.SUBSYSTEMS (§3.4.3.1.1(c))
ATOM.SITE.OCC.FOURIER.PARAM (§3.4.3.3.2(e))	CELL.WAVE.VECTOR (§3.4.3.1.1(d))
ATOM.SITE.OCC.SPECIAL.FUNC (§3.4.3.3.3(b))	CELL.WAVE.VECTORS (§3.4.3.1.1(e))
ATOM.SITE.PHASON (§3.4.3.3.1(b))	DIFFRN group (§3.4.3.1.2)
ATOM.SITE.ROT.FOURIER (§3.4.3.3.2(f))	DIFFRN.REFLN (§3.4.3.1.2(a)) ¶
ATOM.SITE.ROT.FOURIER.PARAM (§3.4.3.3.2(g))	DIFFRN.REFLNS (§3.4.3.1.2(b)) ¶
ATOM.SITE.U.FOURIER (§3.4.3.3.2(h))	DIFFRN.STANDARD.REFLN (§3.4.3.1.2(c)) ¶
ATOM.SITE.U.FOURIER.PARAM (§3.4.3.3.2(i))	EXPTL group (§3.4.3.1.3)
ATOM.SITES.DISPLACE.FOURIER (§3.4.3.3.2(j))	EXPTL.CRYSTAL (§3.4.3.1.3(a)) ¶
ATOM.SITES.MODULATION (§3.4.3.3.2(k))	EXPTL.CRYSTAL.FACE (§3.4.3.1.3(b)) ¶
ATOM.SITES.ROT.FOURIER (§3.4.3.3.2(l))	GEOM group (§3.4.3.3.4)
	GEOM.ANGLE (§3.4.3.3.4(a)) ¶
	GEOM.BOND (§3.4.3.3.4(b)) ¶
	GEOM.CONTACT (§3.4.3.3.4(c)) ¶
	GEOM.TORSION (§3.4.3.3.4(d)) ¶
	REFINE group (§3.4.3.2.1)
	REFINE (§3.4.3.2.1) ¶
	REFLN group (§3.4.3.2.2)
	REFLN (§3.4.3.2.2(a)) ¶
	REFLNS (§3.4.3.2.2(b)) ¶
	SYMMETRY group (§3.4.3.3.5)
	SPACE.GROUP (§3.4.3.3.5(a)) ¶
	SPACE.GROUP.SYMOP (§3.4.3.3.5(b)) ¶

¶ Categories already defined in the core CIF dictionary but containing new items.

‡ A stricter definition than that given in the core CIF dictionary.

The msCIF dictionary exists thanks to the continuous help and support received from I. David Brown and Brian McMahon. The suggestions and criticisms of Sander van Smaalen, Howard Flack, Juan Manuel Pérez-Mato and, especially, Vaclav Petříček (who has implemented a CIF input/output routine in the program *JANA2000*) were important.

### References

- Cervellino, A., Haibach, T. & Steurer, W. (2002). *Structure solution of the basic decagonal Al–Co–Ni phase by the atomic surfaces modelling method*. *Acta Cryst.* **B58**, 8–33.
- Chapuis, G., Farkas-Jahnke, M., Pérez-Mato, J. M., Senechal, M., Steurer, W., Janot, C., Pandey, D. & Yamamoto, A. (1997). *Checklist for the description of incommensurate modulated crystal structures. Report of the International Union of Crystallography Commission on Aperiodic Crystals*. *Acta Cryst.* **A53**, 95–100.
- Elcoro, L., Pérez-Mato, J. M., Darriet, J. & El Abed, A. (2003). *Super-space description of trigonal and orthorhombic  $A_{1+x}A'_xB_{1-x}O_3$  compounds as modulated layered structures; application to the refinement of trigonal  $Sr_6Rh_5O_{15}$* . *Acta Cryst.* **B59**, 217–233.
- Haibach, T., Cervellino, A., Estermann, M. A. & Steurer, W. (2000). *X-ray structure determination of quasicrystals – limits and potentiality*. *Z. Kristallogr.* **215**, 569–583.