

3.4. CLASSIFICATION AND USE OF MODULATED AND COMPOSITE STRUCTURES DATA

modulation of the atom-site properties. They fall naturally into families describing the modulation of atomic displacement, of site occupation or of thermal parameters.

New data items are added to several categories in the core CIF dictionary that describe molecular or packing geometry. There are also new data items to describe superspace-group symmetry.

3.4.3.3.1. *Atom sites*

Data items in these categories are as follows:

(a) ATOM_SITE

```
_atom_site_displace_modulation_flag
_atom_site_occ_modulation_flag
_atom_site_subsystem_code
  → _cell_subsystem_code
_atom_site_U_modulation_flag
```

(b) ATOM_SITE_PHASON

```
• _atom_site_phason_atom_site_label
  → _atom_site_label
_atom_site_phason_coeff
_atom_site_phason_formula
```

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item.

The ATOM_SITE category is extended in the msCIF dictionary by the addition of a small number of items that may appear in the main looped list of atom-site information (see Section 3.2.4.1.1). The **_flag* items indicate whether each individual atom site has been modelled through modulation of atomic displacement, site occupation or thermal parameters. In each case, the default value of the item is *no*, so that any or all of the flags may be omitted when that particular type of modulation has not been applied to the structural model.

_atom_site_subsystem_code identifies the cell subsystem to which the atom site must be assigned in the description of composite structures. Each value of *_atom_site_subsystem_code* must match one of the values of *_cell_subsystem_code* in the overall description of the subsystems defined for a composite.

The ATOM_SITE_PHASON category allow details of an atom-dependent phason correction, as implemented in *JANA2000*, to be given. The use of these phason corrections is discouraged.

3.4.3.3.2. *Modulation functions as Fourier series*

Data items in these categories are as follows:

(a) ATOM_SITE_DISPLACE_FOURIER

```
• _atom_site_displace_Fourier_id
  _atom_site_displace_Fourier_atom_site_label
  → _atom_site_label
  _atom_site_displace_Fourier_axis
  _atom_site_displace_Fourier_wave_vector_seq_id
  → _atom_site_Fourier_wave_vector_seq_id
```

(b) ATOM_SITE_DISPLACE_FOURIER_PARAM

```
• _atom_site_displace_Fourier_param_id
  → _atom_site_displace_Fourier_id
  _atom_site_displace_Fourier_param_cos
  _atom_site_displace_Fourier_param_modulus
  _atom_site_displace_Fourier_param_phase
  _atom_site_displace_Fourier_param_sin
```

(c) ATOM_SITE_FOURIER_WAVE_VECTOR

```
_atom_site_Fourier_wave_vector_description
_atom_site_Fourier_wave_vector_seq_id
_atom_site_Fourier_wave_vector_x
_atom_site_Fourier_wave_vector_y
_atom_site_Fourier_wave_vector_z
```

(d) ATOM_SITE_OCC_FOURIER

```
• _atom_site_occ_Fourier_id
  _atom_site_occ_Fourier_atom_site_label
  → _atom_site_label
```

```
_atom_site_occ_Fourier_wave_vector_seq_id
  → _atom_site_Fourier_wave_vector_seq_id
```

(e) ATOM_SITE_OCC_FOURIER_PARAM

```
• _atom_site_occ_Fourier_param_id
  → _atom_site_occ_Fourier_id
  _atom_site_occ_Fourier_param_cos
  _atom_site_occ_Fourier_param_modulus
  _atom_site_occ_Fourier_param_phase
  _atom_site_occ_Fourier_param_sin
```

(f) ATOM_SITE_ROT_FOURIER

```
• _atom_site_rot_Fourier_id
  _atom_site_rot_Fourier_atom_site_label
  → _atom_site_label
  _atom_site_rot_Fourier_axis
  _atom_site_rot_Fourier_wave_vector_seq_id
  → _atom_site_Fourier_wave_vector_seq_id
```

(g) ATOM_SITE_ROT_FOURIER_PARAM

```
• _atom_site_rot_Fourier_param_id
  → _atom_site_rot_Fourier_id
  _atom_site_rot_Fourier_param_cos
  _atom_site_rot_Fourier_param_modulus
  _atom_site_rot_Fourier_param_phase
  _atom_site_rot_Fourier_param_sin
```

(h) ATOM_SITE_U_FOURIER

```
• _atom_site_U_Fourier_id
  _atom_site_U_Fourier_atom_site_label
  → _atom_site_label
  _atom_site_U_Fourier_tens_elem
  _atom_site_U_Fourier_wave_vector_seq_id
  → _atom_site_Fourier_wave_vector_seq_id
```

(i) ATOM_SITE_U_FOURIER_PARAM

```
• _atom_site_U_Fourier_param_id
  → _atom_site_U_Fourier_id
  _atom_site_U_Fourier_param_cos
  _atom_site_U_Fourier_param_modulus
  _atom_site_U_Fourier_param_phase
  _atom_site_U_Fourier_param_sin
```

(j) ATOM_SITES_DISPLACE_FOURIER

```
_atom_sites_displace_Fourier_axes_description
```

(k) ATOM_SITES_MODULATION

```
_atom_sites_modulation_global_phase_t_1
_atom_sites_modulation_global_phase_t_2
_atom_sites_modulation_global_phase_t_3
_atom_sites_modulation_global_phase_t_4
_atom_sites_modulation_global_phase_t_5
_atom_sites_modulation_global_phase_t_6
_atom_sites_modulation_global_phase_t_7
_atom_sites_modulation_global_phase_t_8
```

(l) ATOM_SITES_ROT_FOURIER

```
_atom_sites_rot_Fourier_axes_description
```

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item.

It is common to represent a modulated structure using a reference periodic structure on which are superimposed atomic modulation functions expanded as Fourier series. (A full discussion of this is given in Section 3.4.4.3.) The msCIF dictionary provides separate categories for listing the modulated parameters that apply to atom positions, site occupancies and thermal parameters. The structuring of the data items within each of these categories follows a similar pattern.

For example, consider the modulation of the atomic displacements. The ATOM_SITE_DISPLACE_FOURIER category allows a listing of the axis along which the displacement occurs (**_axis*) and the wave vectors contributing to that displacement component (**_wave_vector_seq_id*) for each relevant atom site (labelled by *_atom_site_displace_Fourier_atom_site_label*).

3. CIF DATA DEFINITION AND CLASSIFICATION

`*_wave_vector_seq_id` is a pointer to the description of the separate modulation wave vectors and must match one of the identifiers `_atom_site_Fourier_wave_vector_seq_id` listed separately in the `ATOM_SITE_FOURIER_WAVE_VECTOR` category. Likewise, the `*_atom_site_label` data item must match a value of `_atom_site_label` in the main list of atom positions. This is how the modulation is linked to the atom list. The item `_atom_site_displace_Fourier_id` is the formal key for the `ATOM_SITE_DISPLACE_FOURIER` category. It is used to locate the matching Fourier coefficients in the `ATOM_SITE_DISPLACE_FOURIER_PARAM` category. The coefficients may be reported in a sine-cosine (`_atom_site_displace_Fourier_param_sin`, `*_cos`) or modulus-argument (`*_mod`, `*_phase`) representation.

Where a group of atoms is treated as a rigid group, the categories above describe only the translational part of the positional distortion. `ATOM_SITE_ROT_FOURIER` and `ATOM_SITE_ROT_FOURIER_PARAM` are used to describe the rotational components.

`ATOM_SITE_OCC_FOURIER` and `ATOM_SITE_U_FOURIER`, and their associated `*_PARAM` categories, are the analogous categories for the modulation of site occupation and thermal parameters.

All the categories above describe the properties of individual atom sites. Larger-scale descriptions of the displacive modulation or of the rotational component of a rigid group are covered by the categories `ATOM_SITES_DISPLACE_FOURIER` and `ATOM_SITES_ROT_FOURIER`, each of which at present contains one descriptive data item.

The `ATOM_SITES_MODULATION` category contains data items describing the initial phases of the modulation waves, which are essential for determining the space group of the commensurate superstructure. More details are given in the dictionary.

3.4.3.3.3. Special modulation functions

Data items in these categories are as follows:

(a) `ATOM_SITE_DISPLACE_SPECIAL_FUNC`

- `_atom_site_displace_special_func_atom_site_label`
→ `_atom_site_label`
`_atom_site_displace_special_func_sawtooth_ax`
`_atom_site_displace_special_func_sawtooth_ay`
`_atom_site_displace_special_func_sawtooth_az`
`_atom_site_displace_special_func_sawtooth_c`
`_atom_site_displace_special_func_sawtooth_w`

(b) `ATOM_SITE_OCC_SPECIAL_FUNC`

- `_atom_site_occ_special_func_atom_site_label`
→ `_atom_site_label`
`_atom_site_occ_special_func_crenel_c`
`_atom_site_occ_special_func_crenel_w`

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item.

Several data items cover modulation functions that are not expressed as Fourier expansions. The examples in the current msCIF dictionary are restricted to the one-dimensional modulations (sawtooth displacive and occupational crenel functions) implemented in the program *JANA2000* (see Section 3.4.2).

3.4.3.3.4. Molecular or packing geometry

New data items in these categories are as follows:

(a) `GEOM_ANGLE`

- `_geom_angle_av`
`_geom_angle_max`
`_geom_angle_min`
`_geom_angle_site_ssg_symmetry_1`
`_geom_angle_site_ssg_symmetry_2`
`_geom_angle_site_ssg_symmetry_3`

(b) `GEOM_BOND`

- `_geom_bond_distance_av`
`_geom_bond_distance_max`
`_geom_bond_distance_min`
`_geom_bond_site_ssg_symmetry_1`
`_geom_bond_site_ssg_symmetry_2`

(c) `GEOM_CONTACT`

- `_geom_contact_distance_av`
`_geom_contact_distance_max`
`_geom_contact_distance_min`
`_geom_contact_site_ssg_symmetry_1`
`_geom_contact_site_ssg_symmetry_2`

(d) `GEOM_TORSION`

- `_geom_torsion_av`
`_geom_torsion_max`
`_geom_torsion_min`
`_geom_torsion_site_ssg_symmetry_1`
`_geom_torsion_site_ssg_symmetry_2`
`_geom_torsion_site_ssg_symmetry_3`
`_geom_torsion_site_ssg_symmetry_4`

For each of the geometry categories, there are two groups of extensions. One set covers maximum, minimum and average values of bonds, contact distances, angles and torsion angles. The other extends the symmetry-operation code used in geometry listings in the core CIF dictionary (see Section 3.2.4.3.2) to the higher-dimensional superspace form.

3.4.3.3.5. Symmetry information

New data items in these categories are as follows:

(a) `SPACE_GROUP`

- `_space_group_ssg_IT_number`
`_space_group_ssg_name`
`_space_group_ssg_name_IT`
`_space_group_ssg_name_WJJ`
`_space_group_ssg_WJJ_code`

(b) `SPACE_GROUP_SYMOP`

- `_space_group_symop_ssg_id`
`_space_group_symop_ssg_operation_algebraic`

At present, the msCIF dictionary extends the core CIF dictionary symmetry categories to describe superspace groups for one-dimensional modulated structures in four ways: as the superspace-group number in Janssen *et al.* (2004) (`_space_group_ssg_IT_number`), as the *International Tables* superspace-group symbol (`*_ssg_name_IT`), as one of the notations from de Wolff *et al.* (1981) (`*_ssg_name_WJJ`, `*_ssg_WJJ_code`), or in some other formalism (`*_ssg_name`). At present, superspace-group names for higher dimensions can only be indicated using `_space_group_ssg_name`.

Symmetry operations in the superspace group are specified in the `SPACE_GROUP_SYMOP` category by an obvious extension to the method used in the core dictionary. These items must always be present in a CIF corresponding to a modulated or composite structure.

3.4.3.4. File metadata

The categories modified in the msCIF dictionary to formalize the construction of a multi-block description of modulated or composite structures are as follows:

- AUDIT group
AUDIT ¶
AUDIT_LINK ¶

Categories marked with ¶ are already defined in the core CIF dictionary.