

3.6. CLASSIFICATION AND USE OF MACROMOLECULAR DATA

ATOM_SITES_ALT_GEN category explicitly tie together the alternative conformations that contribute to each ensemble. Finally, the atoms in each alternative conformation are identified in the ATOM_SITE category by the data item `_atom_site.label_alt_id`.

The current version of the mmCIF dictionary cannot be used to describe an NMR structure determination completely. However, an mmCIF can be used to store the multiple models usually used to describe a structure determined by NMR using the data items in these categories.

Example 3.6.7.3 is a simplified version of the example given in the mmCIF dictionary (see Fig. 3.6.7.2).

3.6.7.2. Molecular chemistry

The categories describing molecular chemistry are as follows:

Molecular chemistry in the core CIF dictionary (§3.6.7.2.1)

CHEMICAL group

CHEMICAL
CHEMICAL_CONN_ATOM
CHEMICAL_CONN_BOND
CHEMICAL_FORMULA

Chemical components (§3.6.7.2.2)

CHEM_COMP group

CHEM_COMP
CHEM_COMP_ANGLE
CHEM_COMP_ATOM
CHEM_COMP_BOND
CHEM_COMP_CHIR
CHEM_COMP_CHIR_ATOM
CHEM_COMP_PLANE
CHEM_COMP_PLANE_ATOM
CHEM_COMP_TOR
CHEM_COMP_TOR_VALUE

Chemical links (§3.6.7.2.3)

CHEM_LINK group

CHEM_COMP_LINK
CHEM_LINK
CHEM_LINK_ANGLE
CHEM_LINK_BOND
CHEM_LINK_CHIR
CHEM_LINK_CHIR_ATOM
CHEM_LINK_PLANE
CHEM_LINK_PLANE_ATOM
CHEM_LINK_TOR
CHEM_LINK_TOR_VALUE
ENTITY_LINK

The detailed chemistry of the components of a macromolecular structure can be described using data items in the CHEM_COMP and CHEM_LINK category groups. These mmCIF categories are used in preference to those in the CHEMICAL category group in the core CIF dictionary, as macromolecules are in most cases linked assemblies of a limited number of monomers and so they are most efficiently described by defining the monomers and the links between them, rather than by a formal definition of every bond and angle.

All the categories relevant to molecular chemistry are listed in the summary above; note in particular the presence of the category ENTITY_LINK within the formal CHEM_LINK category group.

3.6.7.2.1. Molecular chemistry in the core CIF dictionary

The data items in these categories are as follows:

(a) CHEMICAL

- `_chemical.entry_id`
→ `_entry.id`

- `_chemical.absolute_configuration`
- `_chemical.compound_source`
- `_chemical.melting_point`
- `_chemical.melting_point_gt`
- `_chemical.melting_point_lt`
- `_chemical.name_common`
- `_chemical.name_mineral`
- `_chemical.name_structure_type`
- `_chemical.name_systematic`
- `_chemical.optical_rotation`
- `_chemical.properties_biological`
- `_chemical.properties_physical`
- + `_chemical.temperature_decomposition`
- `_chemical.temperature_decomposition_gt`
- `_chemical.temperature_decomposition_lt`
- + `_chemical.temperature_sublimation`
- `_chemical.temperature_sublimation_gt`
- `_chemical.temperature_sublimation_lt`

(b) CHEMICAL_CONN_ATOM

- `_chemical_conn_atom.number`
- `_chemical_conn_atom.charge`
- `_chemical_conn_atom.display_x`
- `_chemical_conn_atom.display_y`
- `_chemical_conn_atom.NCA`
- `_chemical_conn_atom.NH`
- `_chemical_conn_atom.type_symbol`

(c) CHEMICAL_CONN_BOND

- `_chemical_conn_bond.atom_1`
- `_chemical_conn_bond.atom_2`
- `_chemical_conn_bond.type`

(d) CHEMICAL_FORMULA

- `_chemical_formula.entry_id`
→ `_entry.id`
- `_chemical_formula.analytical`
- `_chemical_formula.iupac`
- `_chemical_formula.moiety`
- `_chemical_formula.structural`
- `_chemical_formula.sum`
- `_chemical_formula.weight`
- `_chemical_formula.weight_meas`

The bullet (•) indicates a category key. Where multiple items within a category are marked with a bullet, they must be taken together to form a compound key. The arrow (→) is a reference to a parent data item. Items in italics have aliases in the core CIF dictionary formed by changing the full stop (.) to an underscore (_). Data items marked with a plus (+) have companion data names for the standard uncertainty in the reported value, formed by appending the string `_esd` to the data name listed.

Descriptions of molecular chemistry in an mmCIF are normally made using data items in the CHEM_COMP and CHEM_LINK category groups. The CHEMICAL category group is retained in the mmCIF dictionary solely for consistency with the core CIF dictionary and Section 3.2.4.2 may be consulted for details.

Two of the categories in this group, CHEMICAL_CONN_ATOM and CHEMICAL_CONN_BOND, have existing category keys in the core dictionary. The formal keys `_chemical.entry_id` and `_chemical_formula.entry_id` have been added to CHEMICAL and CHEMICAL_FORMULA, respectively, to provide the category keys required by the DDL2 data model.

It is emphasized that these items will not appear in the description of a macromolecular structure, but they are retained to allow the representation of small-molecule or inorganic structures in the DDL2 formalism of mmCIF.

3.6.7.2.2. Chemical components

Data items in these categories are as follows:

(a) CHEM_COMP

- `_chem_comp.id`
- `_chem_comp.formula`
- `_chem_comp.formula_weight`
- `_chem_comp.model_details`
- `_chem_comp.model_eref`