```
data_mmcif_std.dic

    _dictionary.title           mmcif_std.dic
    _dictionary.version         2.0.09
    _dictionary.datablock_id    mmcif_std.dic
                          (a)

     loop_
         _dictionary_history.version
         _dictionary_history.update
         _dictionary_history.revision    .    .    .
                          (b)

     loop_
    _sub_category.id
    _sub_category.description     .    .

     loop_
    _category_group_list.id
    _category_group_list.parent_id
    _category_group_list.description    .    .    .
                          (c)

    loop_
    _item_type_list.code
    _item_type_list.primitive_code
    _item_type_list.construct
    _item_type_list.detail

    loop_
    _item_units_list.code
    _item_units_list.detail    .    .

     loop_
    _item_units_conversion.from_code
    _item_units_conversion.to_code
    _item_units_conversion.operator
    _item_units_conversion.factor    .    .    .    .
                          (d)


  save_CATEGORY_A    .    .    .    save_
    save__category_a.item_1    .    .    .    save_
    save__category_a.item_2    .    .    .    save_
    save__category_a.item_3    .    .    .    save_

  save_CATEGORY_B    .    .    .    save_
    save__category_b.item_1    .    .    .    save_
    save__category_b.item_2    .    .    .    save_
                          (e)
```

Fig. 3.1.6.1. Schematic structure of the macromolecular CIF dictionary. (*a*) Dictionary identifiers. (*b*) Dictionary history. (*c*) Subcategory and category group listings. (*d*) Data types, units descriptions and conversion tables. (*e*) Multiple category and item definition blocks.

items are defined in the same data block and are understood to share the common attributes itemized in that data block).

Within DDL2, there are mechanisms for more formal and machine-parsable statements of relationships. The `_sub_category.id` attribute is a label shared by several data items within a category that are related in a specific way described by the associated `_sub_category.description` attribute. The relationships may be rather general, such as elements of a matrix; or they may be specific physical properties or attributes, such as the collection of axis lengths of a unit cell. The dictionary should list all such labels that occur within its included data definition blocks. Example 3.1.6.2 is an extract from the macromolecular dictionary.

3.1.6.3. Category groupings

In the DDL2 data model, a *category* of data corresponds to a set of related data items that may be stored in a single relational database table. A number of such tables may collectively describe the complete properties of some physical object. This is expressed formally by assigning the same label (`_category_group.id`) to the relevant categories. While relationships between categories are implied in DDL1 dictionaries by the hierarchical structure of the names of data items, in DDL2 dictionaries the relationships are formally stated.

For subcategories, the category-group relationships present in the dictionary are listed in a separate looped list. Example 3.1.6.3 is an extract from the macromolecular dictionary. The `inclusive_group` entry shows the common parentage of all categories (and ultimately all data items) in the dictionary.

Example 3.1.6.1. *DDL2 dictionary identification entries.*

```
data_mmcif_std.dic

    _dictionary.title           mmcif_std.dic
    _dictionary.version         2.0.09
    _dictionary.datablock_id    mmcif_std.dic

     loop_
    _dictionary_history.version
    _dictionary_history.update
    _dictionary_history.revision
   0.1.1  1993-02-11
;  Highlighted all notes with # %%%%% surrounds.
;
   .  .  .
```

Example 3.1.6.2. *DDL2 subcategories defined in the mmCIF dictionary.*

```
loop_
    _sub_category.id
    _sub_category.description
    'fractional_coordinate'
; The collection of x, y, and z components of a
  position specified with reference to unit cell
  directions.
;
   'matrix'
; The collection of elements of a matrix.
;
   'miller_index'
; The collection of h, k, and l components of the
  Miller index of a reflection.
;
   'cell_length'
; The collection of a, b, and c axis lengths of a
  unit cell.
;
   'mm_atom_site_label'
; The collection of alt id, asym id, atom id, comp id
  and seq id components of the label for a
  macromolecular atom site.
;
```

3.1.6.4. Category definitions

In the DDL2 formalism, a category of data items may be mapped to a relational table. The dictionary entry for a category includes the name of the category (an identifying label which is referenced by the `_item.category_id` attribute of each component data item) and a list of the category groups of which it may be considered a member. The category *key* is explicitly specified – that is, the data item (or group of items) that uniquely identifies an individual row in a table of data of that category.

Where a category encompasses a set of data items that are not normally specified in a looped list, the category may nevertheless be taken to represent a degenerate table with a single row, and therefore there is still a category key. For degenerate categories the key value is often set equal to the name of the parent data block.

Example 3.1.6.3. *Category groups in a DDL2 dictionary.*

```
loop_
  _category_group_list.id
  _category_group_list.parent_id
  _category_group_list.description
    'inclusive_group'    .
; Categories that belong to the macromolecular
  dictionary.
;
    'atom_group'
    'inclusive_group'
; Categories that describe the properties of atoms.
;
    'audit_group'
    'inclusive_group'
; Categories that describe dictionary maintenance and
  identification.
;
    'cell_group'
    'inclusive_group'
; Categories that describe the unit cell.
;
```

Example 3.1.6.4. *A category description in a DDL2 dictionary.*

```
save_EXPTL
    _category.description
;   Data items in the EXPTL category record details
    about the experimental work prior to the
    intensity measurements and details about the
    absorption-correction technique employed.
;
    _category.id                    exptl
    _category.mandatory_code        no
    _category_key.name            '_exptl.entry_id'
     loop_
    _category_group.id             'inclusive_group'
                                   'exptl_group'
     loop_
    _category_examples.detail
    _category_examples.case
# - - - - - - - - - - - - - - - - - - - - - - - - - -
;  Example 1 - based on laboratory records for
   Yb(S-C5H4N)2 (THF)4
;
; _exptl.entry_id                     datablock1
  _exptl.absorpt_coefficient_mu       1.22
  _exptl.absorpt_correction_T_max     0.896
  _exptl.absorpt_correction_T_min     0.802
  _exptl.absorpt_correction_type      integration
  _exptl.absorpt_process_details
  ; Gaussian grid method from SHELX76
    Sheldrick, G. M., "SHELX-76: structure
    determination and refinement program",
    Cambridge University, UK, 1976
  ;
  _exptl.crystals_number              1
  _exptl.details
  ; Enraf-Nonius LT2 liquid nitrogen
    variable-temperature device used
  ;
  _exptl.method    'single-crystal x-ray diffraction'
  _exptl.method_details
  ; graphite monochromatized Cu K(alpha) fixed tube
    and Enraf-Nonius CAD4 diffractometer used
  ;
;
# - - - - - - - - - - - - - - - - - - - - - - - - - -
save_
```

Example 3.1.6.5. *A DDL2 category with a composite key.*

```
save_GEOM_BOND
    _category.description
;   Data items in the GEOM_BOND category record
    details about the bond lengths as calculated
    from the contents of the ATOM, CELL and
    SYMMETRY data.
;
    _category.id                    geom_bond
    _category.mandatory_code        no
     loop_
    _category_key.name     '_geom_bond.atom_site_id_1'
                           '_geom_bond.atom_site_id_2'
                           '_geom_bond.site_symmetry_1'
                           '_geom_bond.site_symmetry_2'
     loop_
    _category_group.id             'inclusive_group'
                                   'geom_group'
     loop_
    _category_examples.detail
    _category_examples.case
# - - - - - - - - - - - - - - - - - - - - - - - - - -
;  Example 1 - based on data set TOZ of Willis,
   Beckwith & Tozer [Acta Cryst. (1991), C47,
   2276-2277].
;
;    loop_
     _geom_bond.atom_site_id_1
     _geom_bond.atom_site_id_2
     _geom_bond.dist
     _geom_bond.dist_esd
     _geom_bond.site_symmetry_1
     _geom_bond.site_symmetry_2
     _geom_bond.publ_flag
      O1  C2   1.342  0.004  1_555  1_555  yes
      O1  C5   1.439  0.003  1_555  1_555  yes
      C2  C3   1.512  0.004  1_555  1_555  yes
      C2  O21  1.199  0.004  1_555  1_555  yes
      C3  N4   1.465  0.003  1_555  1_555  yes
      C3  C31  1.537  0.004  1_555  1_555  yes
      C3  H3   1.00   0.03   1_555  1_555  ?
      N4  C5   1.472  0.003  1_555  1_555  yes
    # - - - - data truncated for brevity - - - -
;
# - - - - - - - - - - - - - - - - - - - - - - - - - -
save_
```

Example 3.1.6.4 shows a category of non-looped core data items. It may be compared with the DDL1 version in Example 3.1.5.2.

For categories of looped items (those normally presented in a table of values) it is sometimes appropriate to have as the category key a data item that has the sole function of indexing unique table rows. However, it is also often the case that a composite key is formed from existing data items, and in these cases the category definition must loop the components of the key, as in Example 3.1.6.5 from the macromolecular dictionary definition of the GEOM_BOND category.

It must be remembered that, in practice, data files may lack some of the items required to determine the category key formally. For example, in the data set given in the GEOM_BOND example here, it is possible that the **_geom_bond.site_symmetry_** items may be absent because the listing is for a single connected molecule within an asymmetric unit. Robust parsing software must construct data keys by assigning NULL or other suitable default values to the missing key components.

Careful inspection of corresponding definitions in the DDL1 and DDL2 versions of core data items will demonstrate that the explicit category key specification in DDL2 dictionaries may be deduced within DDL1 dictionaries from the appropriate **_list_reference**, **_list_mandatory** and **_list_uniqueness** attributes of data-item definitions within a category (see also Section 2.5.6.4).

3.1.6.5. Data-item definitions

The bulk of a DDL2 data dictionary comprises the save frames that include descriptions of the meaning and properties of individual data names.

Unlike DDL1 dictionaries, where the definitions of several data names may be contained in a single data block (most commonly for a set of items that form a logical irreducible set), save frames in