3.1. GENERAL CONSIDERATIONS WHEN DEFINING A CIF DATA ITEM

Example 3.1.5.1. *A DDL1 dictionary identification block.*

```
 data_on_this_dictionary
    _dictionary_name            cif_core.dic
    _dictionary_version         2.3.1
    _dictionary_update          2005-06-27
    _dictionary_history
;  1991-05-27  Created from CIF Dictionary text. SRH
   1991-05-30  Validated with CYCLOPS & CIF ms. SRH
   ...
;
```

Example 3.1.5.2. *A category description in a DDL1 dictionary.*

```
data_exptl_[]
    _name                       '_exptl_[]'
    _category                   category_overview
    _type                       null
    loop_  _example
    _example_detail
# - - - - - - - - - - - - - - - - - - - - - - - - -
;   _exptl_absorpt_coefficient_mu    0.962
    _exptl_absorpt_correction_type   psi-scan
    _exptl_absorpt_process_details
            'North, Phillips & Mathews (1968)'
    _exptl_absorpt_correction_T_min  0.929
    _exptl_absorpt_correction_T_max  0.997
;
;   Example 1 - based on a paper by Steiner [Acta
            Cryst. (1996), C52, 2554-2556].
;
# - - - - - - - - - - - - - - - - - - - - - - - - -
    _definition
;     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.
;
```

---

be seen that there is no sure way of working out the category from the complete data name except by referring to its `_category` attribute in the associated dictionary. This differs from the DDL2 convention of including an explicit separator (a full stop) between the category name and the remainder of a data name.

While it is not mandatory that a data name should incorporate its category name as a leading component, authors are strongly encouraged to adopt this convention. A small number of core data items that did not conform to this convention have been deprecated in later releases of the core dictionary. However, in the powder dictionary the convention has been broken so that one can present data sets separately or merge them together. In this dictionary, some data names beginning with the strings `_pd_calc`, `_pd_meas` and `_pd_proc` all belong formally to the category PD_DATA. This allows calculated data values to be tabulated with raw and processed measurements if this is useful.

One other case where a data name does not begin with its associated category name is that of the pseudo data names such as `_exptl_[]` that appear in the dictionary to describe the purpose of a category (Section 3.1.5.3). Such data names are always assigned the category CATEGORY_OVERVIEW and are further differentiated from other data names by having a data type of 'null'.

3.1.5.1. The dictionary identification block

As mentioned above, the dictionary file must contain information that unambiguously states its identity and version. In DDL1-based CIF dictionaries, this is achieved by itemizing the full set of dictionary attributes (see Section 2.5.6.5) within a data block named `data_on_this_dictionary`, as in Example 3.1.5.1 from the core dictionary.

3.1.5.2. Irreducible sets of data items

In general, a dictionary data block defines a single data item. However, there are instances where several related data names are defined in the same data block. Sometimes this has been done for convenience, to produce a compact listing of similar data names that have common attributes and whose small differences in meaning can best be expressed by a single definition. Such groupings are discouraged, except where they represent components of a larger entity that has no sensible meaning in the absence of any of the components. For example, the data block `data_refln_index_` defines the three data items `_refln_index_h`, `_refln_index_k` and `_refln_index_l` that represent the Miller indices of a reflection. All three indices must have a value in order to specify a reflection and so each has no meaning in isolation.

Note that there is no formal method of expressing this close relationship within DDL1 except by grouping the definitions in the same data block in this way. In DDL2 dictionaries, it is common to assign the components of an irreducible set to a specific subcategory.

3.1.5.3. Category descriptions

As discussed above, categories in DDL1 are intended as 'natural groupings' of data items. To document the purpose of a category within a dictionary, 'pseudo' data names are used. All pseudo data names are assigned a `_category` attribute of category_overview and have an associated `_type` value of 'null'. They are also named by convention as `_category_name_[dictionarycode]`, for example `_pd_data_[pd]` for the description of the PD_DATA category in the powder dictionary (indicated by the code 'pd' in square brackets). For the core dictionary, *dictionarycode* is not given, resulting in names like `_exptl_[]` to describe the EXPTL category.

Example 3.1.5.2 is a slightly edited extract from the core dictionary showing how a data block for a category description is composed, including the presence of an example.

Note that the *dictionarycode* extension allows a dictionary to include comments on items that it defines in a category already established in the core dictionary. For example, the modulated structures dictionary includes the category overview item `_audit_link_[ms]`. This describes the convention adopted to express the relationship between data blocks in a modulated structures data file using the `_audit_link_` data names already defined in the core dictionary.

3.1.5.4. Data-item definitions

The data blocks described in Sections 3.1.5.1 and 3.1.5.3 are used to identify the dictionary and to describe the nature and purpose of a category. The remaining data blocks in a dictionary provide the attributes of data values in a form suitable for machine extraction and validation. The following examples show how this is done for various types of data.

3.1.5.4.1. *Definitions of single quantities*

Example 3.1.5.3 is the core dictionary definition of the data name for the ambient temperature during the experiment. Because this is a single (non-looped) value, the relevant data name is one among several discrete items in the DIFFRN category. No further description of its relationship to other data items is required.

The type of the associated data value (*numb* for numerical) is specified, together with any constraint on its legal value. The range

references