3.7. Classification and use of image data

By H. J. Bernstein

3.7.1. Introduction

This chapter describes the categories and organization of data items defined in the CBF/imgCIF dictionary. The classification of image data applies to both Crystallographic Binary File (CBF) and Image-supporting Crystallographic Information File (imgCIF) representations. An introduction to CBF data and construction is given in Chapter 2.3. Full details of the CBF/imgCIF dictionary are given in Chapter 4.6.

The main reason for introducing the new items defined in the CBF/imgCIF dictionary was to extend the mmCIF dictionary (Chapter 3.6) to allow the storage of synchrotron diffraction images. However, these items are also important in other fields that use binary image data, including the publication of articles, the creation of web pages and the production of movies.

Data categories in the CBF/imgCIF dictionary can describe one-, two- and three-dimensional array detectors that output data organized by time and/or wavelength. The categories defined at present support modular data that can be extended for future applications without having to make fundamental structural changes. For example, it is anticipated that additional data items will be needed soon to allow higher-dimensional data representations and more complex data structures; these should be accommodated easily.

The CBF/imgCIF dictionary consists of three groups of categories of data items: the ARRAY_DATA group, the AXIS group and the DIFFRN group (Table 3.7.1.1). All fall within the 'Experimental measurements' classification of Table 3.1.10.1. The DIFFRN group already exists in the mmCIF dictionary (Section 3.6.5.2; see also Section 3.2.2.2) and describes the diffraction data and their measurement. Definitions in the CBF/imgCIF dictionary extend and in some cases restate the definitions in the mmCIF dictionary.

The data categories defined in the CBF/imgCIF dictionary are described in this chapter. Table 3.7.1.1 lists the formal category groups declared in the dictionary and the sections of this chapter in which they are discussed. Each section is divided into subsections describing a single category or a small set of closely related categories. Within each subsection, the data names within the relevant categories are listed. Category keys, pointers to parent data items and aliases to data items in the mmCIF dictionary are indicated.

The data collected in an experiment are organized into scans. Each scan consists of one or more frames. Each frame consists of one or more data arrays. The logical data in the data arrays need to be described in terms of physical arrays of image elements. The axes of the laboratory coordinate system needed to describe the physical positions of the image elements and the positioning of the specimen are given in the AXIS category. The axes used for the positioning systems for the specimen and the detector are constructed in the same laboratory coordinate system.

Table 3.7.1.1. *Category groups defined in the CBF/imgCIF dictionary*

Section	Category group	Subject covered
Experimental measurements		
3.7.2	ARRAY_DATA	Binary image data
3.7.3	AXIS	Axes required to specify the data collection
3.7.4	DIFFRN	Diffraction experiment

The DIFFRN_DETECTOR_AXIS category relates detector elements to axes. The DIFFRN_MEASUREMENT_AXIS category relates goniometers to axes. The DIFFRN_SCAN_AXIS and DIFFRN_SCAN_FRAME_AXIS categories relate scans to overall axis settings and individual frames to frame-by-frame axis settings, respectively.

The organization of the data in the collected arrays of data is given in the ARRAY_STRUCTURE_LIST category and the physical settings of axes for the centres of pixels that correspond to data points are given in the ARRAY_STRUCTURE_LIST_AXIS category.

3.7.2. Binary image data

The six categories that collectively define the relationship between the sequences of octets in arrays of binary data and the information in the images those octets represent are as follows:

ARRAY_DATA group
The image data (§3.7.2.1)
    ARRAY_DATA
Array elements (§3.7.2.2)
    ARRAY_ELEMENT_SIZE
Intensities (§3.7.2.3)
    ARRAY_INTENSITIES
Organization and encoding of array data (§3.7.2.4)
    ARRAY_STRUCTURE
    ARRAY_STRUCTURE_LIST
    ARRAY_STRUCTURE_LIST_AXIS

3.7.2.1. The image data

Data items in this category are as follows:

ARRAY_DATA
- `_array_data.array_id`
  → `_array_structure.id`
- `_array_data.binary_id`
  `_array_data.data`

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

Each value of the `_array_data.data` data item is a sequence of octets representing a binary image. `_array_data.array_id` and `_array_data.binary_id`, taken together, uniquely identify each image. The value of `_array_data.array_id` is a pointer to `_array_structure.id` to provide the relationship between the sequence of octets and the logical structure of the image. Since multiple images may have the same logical structure, the purpose of `_array_data.binary_id` is to ensure that each image has a unique identifier.

Affiliation: Herbert J. Bernstein, Department of Mathematics and Computer Science, Kramer Science Center, Dowling College, Idle Hour Blvd, Oakdale, NY 11769, USA.

references