3.3. Classification and use of powder diffraction data

BY B. H. TOBY

3.3.1. Introduction

Powder diffraction is used for many purposes. One important use is the determination of crystallographic models (crystal structures), particularly when single crystals are not available. Other common uses of powder diffraction include the identification of unknown materials; determining the amounts of different crystalline phases in a mixture; studies of phase transitions; and measuring changes in lattice constants with composition, pressure or temperature. Residual stress measurements are also frequently made using powder diffraction. Measurements of the degree of preferred orientation of grains in a processed material (texture or pole-figure measurements) are also carried out using powder diffraction; these are useful as they can be used to relate the engineering properties of a material to its processing conditions.

A wide range of instrumentation is used for powder diffraction. Synchrotron, sealed-tube and rotating-anode X-ray sources and both spallation and reactor neutron sources are used. When monochromatic radiation is used, diffraction intensities are measured as a function of the detector setting angle, 2θ. When polychromatic radiation is used, however, energy-dispersive detectors at fixed angles are used. For a pulsed neutron source, the neutron time-of-flight (TOF) is measured to provide energy-dispersive detection. Energy-dispersive X-ray diffraction is not common, but is used for certain specialized measurements such as *in situ* high-pressure studies. Most laboratory powder diffractometers have a single discrete detector, but a few use position-sensitive detectors (PSDs). Multiple detectors or PSDs are common for specialist instruments, for example at large national user facilities.

Historically, X-ray sensitive film in a diffraction camera was used to measure the intensities. At present, diffractometers are much more common than cameras, but film is an inexpensive method of area detection and cameras continue to have some advantages over diffractometers. X-ray film, however, has a non-linear sensitivity and a limited dynamic range, so electronic area detectors and image plates are replacing film for many applications.

3.3.2. Dictionary design considerations

The many applications of powder diffraction and the wide range of different instruments used dictate the design of several aspects of the powder diffraction CIF (pdCIF) data dictionary. In contrast to single-crystal measurements, where observations are reduced to instrument-independent structure factors, in powder diffraction it is not possible to interpret the data without a detailed knowledge of the type of instrument used. Another difference is that while single-crystal measurements are rarely made for any purpose other than structure determination, powder diffraction data have many non-crystallographic applications (*e.g.* identifying the phases in a mixture of unknown composition). In a single-crystal diffraction experiment, one sample yields one crystal structure. In a powder diffraction experiment, however, it is common for the

sample to be a mixture of phases. Analysis of the diffraction data may result in multiple crystal structure models. It is also common to use multiple data sets, for example to fit one crystallographic model to both X-ray and neutron data simultaneously.

The powder dictionary was developed with the following basic objectives in mind. A pdCIF:

(i) should record the data-collection experiment as completely as possible and document any data analysis;

(ii) should be appropriate for the exchange of unprocessed measurements, so that it can be used in national laboratories and other shared facilities;

(iii) may contain more than one data set and/or crystal structure through the use of multiple data blocks;

(iv) may accommodate references between data blocks; and

(v) should recognize and accommodate data from as many different types of instruments as possible.

The CIF syntax, as opposed to the STAR File syntax, is not well suited to handling large multi-dimensional data sets. For some two-dimensional image formats, this deficiency was addressed by the development of imgCIF (see Chapters 2.3 and 3.7). It is also true that the CIF syntax is not well suited to storing unprocessed powder-diffraction measurements from the many instruments that use area detectors (particularly for the case of the three-dimensional data structures needed for modern TOF instruments). Even in these cases, however, diffraction intensities are commonly reduced to simpler representations, such as might be input to a Rietveld refinement program. The pdCIF definitions are intended for use with these reduced data sets.

3.3.3. pdCIF dictionary sections

At the time when the powder CIF dictionary was created, it was recommended practice to give data items in other dictionaries names that distinguished them from items in the core CIF dictionary. This recommendation proved cumbersome and was subsequently withdrawn, but pdCIF data items are all named in a way that distinguishes them clearly from core CIF dictionary data items: data names in the pdCIF dictionary all begin with the string `_pd_` and, likewise, the data categories in the pdCIF dictionary all begin with PD_.

The data items in the pdCIF dictionary are named according to the type of information they contain, rather than the way they will be used in a data structure. For example, data items relating to observed data have names beginning with `_pd_meas_`, even though some data items beginning with `_pd_meas_` may be assigned to the same category as items in other parts of the dictionary. The formal category assignments (see Section 3.1.10.1) were dictated by the aim to make pdCIF as versatile and simple to use as possible. As an example, one might want to put both the observed and calculated intensities in a single table (loop), as would be the case for a Rietveld refinement. This requires that the calculated and observed intensity data items are assigned to the same category. Note that this does not require that these items always appear in a single loop; calculated and observed intensity values could appear in

Affiliation: BRIAN H. TOBY, NIST Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, Maryland 20899-8562, USA.

Table 3.3.3.1. *Category groups defined in the powder CIF dictionary*

The groups are listed in the order in which they are described in this chapter.

Section	Category group	Subject covered
(*a*) Experimental measurements		
3.3.4.1	PD_CHAR	Characterization of a sample
3.3.4.1	PD_PREP	Preparation of a sample
3.3.4.2	PD_SPEC	Specimen used in an experiment
3.3.4.3	PD_CALIB	Calibration information
3.3.4.3	PD_INSTR	The experimental instrument
3.3.4.4	PD_MEAS	Raw measurements and instrumental settings
(*b*) Analysis		
3.3.5.1	PD_PROC	Processed settings
3.3.5.2	PD_CALC	Simulated settings
3.3.5.3	PD_PEAK	Diffraction peak table
3.3.5.4	REFLN	Reflection assignments and intensities
(*c*) Atomicity, chemistry and structure		
3.3.6.1	PD_PHASE	Phases present
(*d*) File metadata		
3.3.7.1	PD_BLOCK	Relationships between data blocks
3.3.4.3, 3.3.4.4, 3.3.5.1, 3.3.5.2	PD_DATA	Measured and simulated intensities

separate loops, for example, if the increment between data points differs.

This need to contain diverse items in a common 'looped' list has led the pdCIF dictionary to use category names in a different way from the other CIF dictionaries, in which CIF data items are usually named according to their category. In the pdCIF dictionary, data items that might appear in the loop for diffraction intensities are assigned to the category PD_DATA. Only one data item is named using this category as prefix, `_pd_data_point_id`. Another departure from the convention used in other dictionaries is that several `_pd_refln_*` data names are assigned to the category REFLN so that these items may be included in a loop with `_refln_*` items defined in the core CIF dictionary.

Table 3.3.3.1 summarizes the category groups in the pdCIF dictionary; the individual categories are listed alphabetically in Appendix 3.3.1. The appendix also lists for each category the section of this chapter in which the category is described.

The order in which the categories are discussed follows the scheme of Table 3.1.10.1, so that the contents of the dictionary are summarized under the headings *Experimental measurements* (Section 3.3.4), *Analysis* (Section 3.3.5), *Atomicity, chemistry and structure* (Section 3.3.6) and *File metadata* (Section 3.3.7). The pdCIF dictionary does not contribute any new data items relevant to publication beyond those already in the core CIF dictionary.

The data items in each category are listed below. Category keys, if specified, are listed first and are marked by a bullet (•); the remaining data items in each category are listed alphabetically. Note that the category PD_DATA is discussed in several different sections.

3.3.4. Experimental measurements

The categories in the powder CIF dictionary relating to the crystallographic experiment are as follows:

Characterization and preparation of the sample (§3.3.4.1)
  PD_CHAR group
    PD_CHAR
  PD_PREP group
    PD_PREP

Description of the specimen (§3.3.4.2)
  PD_SPEC group
    PD_SPEC
Instrument calibration and design (§3.3.4.3)
  PD_CALIB group
    PD_CALIB
    PD_CALIBRATION
  PD_INSTR group
    PD_INSTR
Observations and measurement conditions (§3.3.4.4)
  PD_DATA group
    PD_DATA (items beginning with `_pd_meas_*`)
  PD_MEAS group
    PD_MEAS_INFO
    PD_MEAS_METHOD

The pdCIF dictionary differentiates between the terms *sample* and *specimen*. The terms are often treated as interchangeable, but they have quite distinct meanings. The term sample refers to a batch of material, while the term specimen refers to the particular portion of the sample that was used for a measurement. In some cases, the specimen is modified before it is used for data collection. For example, it may be mixed with an internal standard, dried, hydrated or pressed into a pellet.

3.3.4.1. Characterization and preparation of the sample

The data items in these categories are as follows:

(*a*) PD_CHAR
  `_pd_char_atten_coef_mu_calc`
  `_pd_char_atten_coef_mu_obs`
  `_pd_char_colour`
  `_pd_char_particle_morphology`
  `_pd_char_special_details`

(*b*) PD_PREP
  `_pd_prep_conditions`
  `_pd_prep_cool_rate`
  `_pd_prep_pressure`
  `_pd_prep_temperature`

The PD_CHAR data items describe information known about the sample from observation and chemical analysis. For example, a description of the sample morphology can be specified using `_pd_char_particle_morphology`. Note that there are data items in the core dictionary that are appropriate for use with powder diffraction. For example, `_atom_type_analytical_mass_%` can be used for chemical analysis results and `_chemical_melting_point` for the melting point. Several similar data items occur in the pdCIF and core dictionaries. `_exptl_crystal_colour` and `_pd_char_colour` both describe the sample colour, but `_pd_char_colour` is more systematic. Also, `_pd_char_atten_coef_mu_calc` and `_exptl_absorpt_coefficient_mu` describe similar properties, but `_pd_char_atten_coef_mu_calc` is adjusted for the sample packing fraction, so it can be compared with the experimental value, `_pd_char_atten_coef_mu_obs`, when a direct measurement is made.

The PD_PREP data items describe how the sample was collected or prepared. For example, `_pd_prep_pressure` and `_pd_prep_temperature` describe the pressure and temperature used to prepare the sample. Note that these will probably differ from the pressure and temperature conditions at which diffraction measurements are made. Measurement conditions are recorded in `_diffrn_ambient_pressure` and `_diffrn_ambient_temperature`.

3.3.4.2. Description of the specimen used in the experiment

The data items in this category are as follows:

PD_SPEC
```
  _pd_spec_description
  _pd_spec_mount_mode
  _pd_spec_mounting
  _pd_spec_orientation
  _pd_spec_preparation
  _pd_spec_shape
  _pd_spec_size_axial
  _pd_spec_size_equat
  _pd_spec_size_thick
  _pd_spec_special_details
```

The PD_SPEC data items describe the specimen used to measure the diffraction data. The data item **_pd_spec_preparation** describes how the specimen that was used to measure the diffraction data was treated, not how the sample was prepared (PD_PREP) or characterized (PD_CHAR).

The PD_SPEC data items are also used to describe how the specimen was mounted for the diffraction experiment. For example, **_pd_spec_mount_mode** and **_pd_spec_orientation** describe the measurement geometry, while **_pd_spec_shape** and **_pd_spec_size_*** describe the specimen shape and size.

3.3.4.3. Instrument calibration and description

The data items in these categories are as follows:

(*a*) PD_CALIB
```
● _pd_calib_detector_id
  _pd_calib_2theta_offset
  _pd_calib_2theta_off_point
  _pd_calib_2theta_off_min
  _pd_calib_2theta_off_max
  _pd_calib_detector_response
  _pd_calib_std_external_block_id
  _pd_calib_std_external_name
  _pd_calib_std_internal_mass_%
  _pd_calib_std_internal_name
```

(*b*) PD_CALIBRATION
```
  _pd_calibration_conversion_eqn
  _pd_calibration_special_details
```

(*c*) PD_INSTR
```
  _pd_instr_2theta_monochr_pre
  _pd_instr_2theta_monochr_post
  _pd_instr_beam_size_ax
  _pd_instr_beam_size_eq
  _pd_instr_cons_illum_flag
  _pd_instr_cons_illum_len
  _pd_instr_dist_src/mono
  _pd_instr_dist_mono/spec
  _pd_instr_dist_src/spec
  _pd_instr_dist_spec/anal
  _pd_instr_dist_anal/detc
  _pd_instr_dist_spec/detc
  _pd_instr_divg_ax_src/mono
  _pd_instr_divg_ax_mono/spec
  _pd_instr_divg_ax_src/spec
  _pd_instr_divg_ax_spec/anal
  _pd_instr_divg_ax_anal/detc
  _pd_instr_divg_ax_spec/detc
  _pd_instr_divg_eq_src/mono
  _pd_instr_divg_eq_mono/spec
  _pd_instr_divg_eq_src/spec
  _pd_instr_divg_eq_spec/anal
  _pd_instr_divg_eq_anal/detc
  _pd_instr_divg_eq_spec/detc
  _pd_instr_geometry
  _pd_instr_location
  _pd_instr_monochr_pre_spec
  _pd_instr_monochr_post_spec
  _pd_instr_slit_ax_src/mono
  _pd_instr_slit_ax_mono/spec
  _pd_instr_slit_ax_src/spec
  _pd_instr_slit_ax_spec/anal
  _pd_instr_slit_ax_anal/detc
  _pd_instr_slit_ax_spec/detc
  _pd_instr_slit_eq_src/mono
  _pd_instr_slit_eq_mono/spec
  _pd_instr_slit_eq_src/spec
  _pd_instr_slit_eq_spec/anal
  _pd_instr_slit_eq_anal/detc
  _pd_instr_slit_eq_spec/detc
  _pd_instr_soller_ax_src/mono
  _pd_instr_soller_ax_mono/spec
  _pd_instr_soller_ax_src/spec
  _pd_instr_soller_ax_spec/anal
  _pd_instr_soller_ax_anal/detc
  _pd_instr_soller_ax_spec/detc
  _pd_instr_soller_eq_src/mono
  _pd_instr_soller_eq_mono/spec
  _pd_instr_soller_eq_src/spec
  _pd_instr_soller_eq_spec/anal
  _pd_instr_soller_eq_anal/detc
  _pd_instr_soller_eq_spec/detc
  _pd_instr_source_size_ax
  _pd_instr_source_size_eq
  _pd_instr_special_details
```

(*d*) Part of PD_DATA
```
  _pd_instr_var_illum_len
```

The bullet (●) *indicates a category key.*

Calibration information can be placed in the PD_CALIB and PD_CALIBRATION categories. The **_pd_calibration_*** data items are descriptive and will not appear in a loop. The **_pd_calib_*** items may be looped to describe multiple detectors. Correction values for 2θ can be given using the **_pd_calib_2theta_offset** and **_pd_calib_2theta_off_*** data items. A calibration equation can be given using **_pd_calibration_conversion_eqn**. When multiple detectors are used, **_pd_calib_detector_response** is used to indicate the relative performance of each detector. The detector deadtime is specified using the core data item **_diffrn_detector_dtime** (which cannot be looped by detector).

If an internal standard is added to the sample for calibration, this information is specified using **_pd_calib_std_internal_name** to specify the material added and **_pd_calib_std_internal_mass_%** to specify the amount.

When a set of calibration intensities is measured using an external standard, it is possible to include the measurements and the derived results in a separate CIF block. A data block would then use **_pd_calib_std_external_block_id** to link to the block containing the calibration information. See Section 3.3.7 for a discussion of block pointers and block IDs. Note that the use of a unique name for the block ID allows the calibration information to be stored in a separate file, so that the calibration CIF need not be repeated in every CIF that references it.

The PD_INSTR section of the pdCIF dictionary contains terms that describe the instrument used. For example, the instrument or laboratory location is given using **_pd_instr_location**. The instrument type can be indicated using **_pd_instr_geometry**. The instrument geometry can be described in much greater detail using several data items. The geometry is described in terms of four regions of the experiment: radiation source to monochromator (**src/mono**); monochromator to specimen (**mono/spec**); specimen to analyser (**spec/anal**); and analyser to detector (**anal/detc**). If no monochromator is present, the first two regions are combined into radiation source to specimen (**src/spec**). If no analyser is present, the last two regions are combined into specimen to detector (**spec/detc**). Thus two, three or four sets of values describe the dimensions of the instrument and the collimation. For example, **_pd_instr_dist_src/mono** would be used to specify the distance between the radiation source and the monochromator. Alternatively, **_pd_instr_dist_src/spec** would be used to specify the distance between the

radiation source and the specimen if no monochromator was present.

Two methods may be used to describe the slits limiting the divergence in the equatorial plane. The angular divergence allowed by the slits in degrees can be specified using the `_pd_instr_divg_eq_*` data items. Alternatively, the dimensions of the slits in the equatorial direction in millimetres may be specified using `_pd_instr_slit_eq_*`. The dimensions of these slits in the axial direction, *i.e.* the direction perpendicular to the equatorial plane and containing the incident or diffracted beam as appropriate, are specified in millimetres using `_pd_instr_slit_ax_*`. The axial slit lengths, along with the `_pd_instr_dist_*` distances, are useful for estimating the low-angle peak asymmetry (Finger *et al.*, 1994). Note that angular divergence in the axial plane is not a well defined concept for line-focus instruments, but can be specified, where appropriate, using `_pd_instr_divg_ax_*`.

The axial and equatorial directions are shown schematically in Fig. 3.3.4.1. The equatorial plane contains the equatorial direction vectors, as well as the incident beam, the diffracted beam and the scattering vector. The axial plane is perpendicular to the equatorial plane and contains the sample centre, which is the point where the incident and diffracted beams meet. For area-detection instruments, the designations of axial and equatorial directions may be arbitrary.

Soller collimators are described using `_pd_instr_soller_eq_*` data items rather than `_pd_instr_divg_eq_*` data items. It is common practice to specify the Soller collimation in arc-minutes (*e.g.* 30′). However, pdCIF defines these items to have units of degrees, so 30′ would be recorded in the CIF as 0.5. It is not usual to limit the axial divergence, except to reduce low-angle asymmetry, but if this is done, the `_pd_instr_soller_ax_*` data items can be used to define this.

For constant-wavelength instruments, it is common to have a monochromator or filter either before the sample, or after the sample (an analyser), or sometimes both. This is described using `_pd_instr_monochr_pre_spec` and `_pd_instr_monochr_post_spec`. It is rare, but possible, to have both a filter and a monochromator in the same location. The BT-1 neutron powder diffractometer at NIST uses both a Cu(311) monochromator and a graphite filter to attenuate the $\lambda/2$ component. In this case, the two elements would be placed in a loop:

```
loop_   _pd_instr_monochr_pre_spec
  '5 seg. vert. focusing Cu(311) monochromator'
  '3 cm graphite filter'
```

Note that the monochromator and analyser takeoff angles are given using `_pd_instr_2theta_monochr_pre` and `_pd_instr_2theta_monochr_post`. It is useful to record these values for X-ray studies, as they are needed for proper polarization corrections.

In a conventional Bragg–Brentano diffractometer, the divergence slits limit the illumination area at the sample. However, since the φ axis (the sample θ axis) is usually set to bisect the 2θ angle of the detector, the actual length of the area of the sample that is illuminated changes with 2θ. One should choose divergence slits so that the beam does not illuminate areas outside the sample at the lowest diffraction angle used. An alternative method for data collection is to have a divergence slit that opens as 2θ increases, so that a constant area of the sample is illuminated. This is known as a θ-compensating slit. Using a θ-compensating slit provides a better signal-to-noise ratio at larger 2θ values, but means that the diffraction intensities have to be normalized to compensate for the change in illumination. It also introduces greater optical
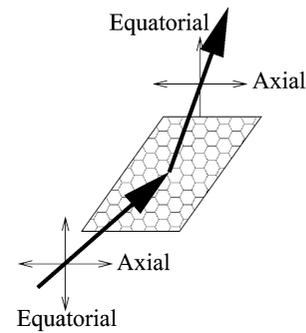

Fig. 3.3.4.1. The axial and equatorial directions in a powder-diffraction experiment.

aberrations with 2θ; a flat plate becomes an increasingly worse approximation to the curved sample geometry in the true Bragg–Brentano geometry.

The use of a variable divergence slit can be recorded in the form:

```
_pd_instr_cons_illum_flag       yes
_pd_instr_cons_illum_len        25.4
```

Note that if `_pd_instr_cons_illum_flag` is not specified, the value is assumed to be no, indicating that a fixed-width divergence slit has been used.

The beam size can be specified in two different ways: as the size at the source, using `_pd_instr_source_size_ax` and `_pd_instr_source_size_eq`, or as the size at the sample position, using `_pd_instr_beam_size_ax` and `_pd_instr_beam_size_eq`. Note that the size of the beam at the sample differs from the illumination length described above except when the sample is perpendicular to the beam. When a variable-divergence slit is in use, the beam size at the sample changes with 2θ, so if this size is known directly, the `_pd_instr_beam_size_eq` data item can be included in the loop containing the diffraction intensities. Similarly, in a constant-divergence instrument, where the illumination length changes with 2θ, the illumination length can be specified in the loop using `_pd_instr_var_illum_len`.

There are also several data items in the core CIF dictionary that should be present in the description of the instrument in a pdCIF. Use `_diffrn_radiation_probe` and `_diffrn_radiation_type` to specify the type of radiation used and `_diffrn_detector_type` to specify the detection type.

3.3.4.4. Observations and measurement conditions

The data items in these categories are as follows:

(*a*) Part of PD_DATA
```
  _pd_data_point_id
  _pd_meas_2theta_fixed
  _pd_meas_2theta_range_min
  _pd_meas_2theta_range_max
  _pd_meas_2theta_range_inc
  _pd_meas_2theta_scan
  _pd_meas_angle_2theta
  _pd_meas_angle_chi
  _pd_meas_angle_omega
  _pd_meas_angle_phi
  _pd_meas_counts_total
  _pd_meas_counts_background
  _pd_meas_counts_container
  _pd_meas_counts_monitor
  _pd_meas_datetime_initiated
  _pd_meas_detector_id
        →  _pd_calib_detector_id
  _pd_meas_point_id
  _pd_meas_position
  _pd_meas_rocking_angle
  _pd_meas_step_count_time
  _pd_meas_time_of_flight
```

(*b*) PD_MEAS_INFO
```
_pd_meas_info_author_address
_pd_meas_info_author_email
_pd_meas_info_author_fax
_pd_meas_info_author_name
_pd_meas_info_author_phone
_pd_meas_intensity_total
_pd_meas_intensity_background
_pd_meas_intensity_container
_pd_meas_intensity_monitor
```

(*c*) PD_MEAS_METHOD
```
_pd_meas_number_of_points
_pd_meas_rocking_axis
_pd_meas_scan_method
_pd_meas_special_details
_pd_meas_units_of_intensity
```

The arrow ($\rightarrow$) is a reference to a parent data item.

The item `_pd_data_point_id` identifies each entry in the list of measured, processed or simulated intensities. It is the only item in the PD_DATA category that actually begins with the string `_pd_data_`, and is included here for convenience. If the list of intensities is split across several distinct loops, the role of this identifier may be adopted by other identifiers, such as `_pd_meas_point_id` in an isolated list of measured intensities.

The `_pd_meas_*` data items contain unprocessed measurements and documentation on the instrumental settings used for the measurements. Note that the choice of the data items used to represent this information is determined by the type of diffraction instrument, as well as how the measurement was conducted. This will be discussed further in Section 3.3.8. However, some `_pd_meas_*` data items are independent of the instrument type. For example, the use of `_pd_meas_datetime_initiated` is good practice, as is use of the `_pd_meas_info_author_*` data items. It is probably good practice to record the number of data points in `_pd_meas_number_of_points` for the benefit of people who might read the CIF, but there is no requirement that this item be present. This means that software should determine the number of data points directly when reading the CIF, rather than relying on the presence of a value for `_pd_meas_number_of_points`.

3.3.5. Analysis

The categories relating to the information derived from the measurements are as follows:

Processed intensities, positions and data processing (§3.3.5.1)
  PD_DATA group
    PD_DATA (items beginning with `_pd_proc_*`)
  PD_PROC group
    PD_PROC_INFO
    PD_PROC_LS
Simulated intensities and their positions (§3.3.5.2)
  PD_DATA group
    PD_DATA (items beginning with `_pd_calc_*`)
  PD_CALC group
    PD_CALC
Diffraction peak table (§3.3.5.3)
  PD_PEAK group
    PD_PEAK
    PD_PEAK_METHOD
Reflection assignments and intensities (§3.3.5.4)
    REFLN

In Rietveld and other studies, processed or simulated intensities are presented alongside measured values. This leads to the presence of both derived and measured values in the same category (PD_DATA). However, the purposes of the data items that

refer to processed and simulated data points are made clear by the way they are named. Overall descriptions of processed and simulated intensity data are covered by the categories PD_PROC_INFO, PD_PROC_LS and PD_CALC. The two categories PD_PEAK and PD_PEAK_METHOD are used to describe lists of peak positions, which would typically be used to search and match powder profiles. Some additional data items relevant to the table of Bragg reflections are defined as additions to the existing REFLN category in the core CIF dictionary.

3.3.5.1. Processed intensities, their positions and processing information

The data items in these categories are as follows:
(*a*) Part of PD_DATA
```
_pd_proc_2theta_corrected
_pd_proc_2theta_range_min
_pd_proc_2theta_range_max
_pd_proc_2theta_range_inc
_pd_proc_d_spacing
_pd_proc_energy_incident
_pd_proc_energy_detection
_pd_proc_intensity_net
_pd_proc_intensity_total
_pd_proc_intensity_bkg_calc
_pd_proc_intensity_bkg_fix
_pd_proc_intensity_incident
_pd_proc_intensity_norm
_pd_proc_point_id
_pd_proc_recip_len_Q
_pd_proc_wavelength
```

(*b*) PD_PROC_INFO
```
_pd_proc_info_author_address
_pd_proc_info_author_email
_pd_proc_info_author_fax
_pd_proc_info_author_name
_pd_proc_info_author_phone
_pd_proc_info_data_reduction
_pd_proc_info_datetime
_pd_proc_info_excluded_regions
_pd_proc_info_special_details
```

(*c*) PD_PROC_LS
```
_pd_proc_ls_background_function
_pd_proc_ls_peak_cutoff
_pd_proc_ls_pref_orient_corr
_pd_proc_ls_prof_R_factor
_pd_proc_ls_prof_wR_factor
_pd_proc_ls_prof_wR_expected
_pd_proc_ls_profile_function
_pd_proc_ls_special_details
_pd_proc_ls_weight
```

The pdCIF dictionary distinguishes between values that are measured directly and values that are derived from these observations. For example, in a constant-wavelength instrument, diffraction intensities are recorded as a function of 2θ. One may derive d-space values from the 2θ values using the value of the wavelength and corrections for the 2θ zero-point error and the sample displacement. One may also derive a new set of data points from the observations, for example by summing adjacent data points when the increment between the data points is much smaller than is warranted by the peak widths. For peak searching and other non-quantitative purposes, the diffraction intensities may be smoothed or otherwise modified. Note that the unprocessed measurement values are retained using the data items `_pd_meas_*`. Since the original measurements are still available, modifications like these do not result in the loss of the original data. In fact, by placing processed values in multiple blocks, a single CIF may contain measurements that have been processed in more than one way.

It is good practice to use the `_pd_proc_info_author_*` and `_pd_proc_info_datetime` data items. It is also a good idea to

describe how the measurements were processed using `_pd_proc_info_data_reduction`.

The `_pd_proc_*` data items in this list may be used to calibrate the 2θ or energy values of the data. These are defined in the items `_pd_proc_2theta_corrected`, `_pd_proc_2theta_range_*`, `_pd_proc_d_spacing`, `_pd_proc_energy_*`, `_pd_proc_recip_len_Q` and `_pd_proc_wavelength`.

When corrections, scaling or other processing, such as averaging or smoothing, are applied to the intensities, the results are stored using the `_pd_proc_intensity_*` data items. Note that if the number of data points does not change, it might be most convenient to include the processed intensities in the same loop as the observed values. This is not always possible, so these items can be placed in a separate loop if there is no longer a one-to-one correspondence between the 2θ or energy positions for the `_pd_proc_intensity_*` values and the `_pd_meas_counts_*` or `_pd_meas_intensity_*` values.

For energy-dispersive measurements, the incident spectrum must be determined for normalization. This can be recorded using `_pd_proc_intensity_incident`. For other types of normalization, `_pd_proc_intensity_norm` should be used.

For full-pattern fitting, there is a series of `_pd_proc_ls_*` data items for recording settings and results. For example, agreement factors can be recorded using the `_pd_proc_ls_prof_*_factor` data items. Some data items may be included in the loop(s) containing the measured or the processed data: `_pd_proc_ls_weight` specifies the weight assigned to each point and `_pd_proc_intensity_bkg_calc` specifies the fitted background. Note that background values are usually generated by extrapolation from fixed values set during the refinement or are determined from a function that is fitted to the observations, and occasionally both are used together. The function that has been fitted can be described using `_pd_proc_ls_background_function`, while fixed points are listed using `_pd_proc_intensity_bkg_fix`. If sections of the pattern are not fitted, this is indicated using `_pd_proc_info_excluded_regions`.

3.3.5.2. Simulated data

The data items in these categories are as follows:
(*a*) Part of PD_DATA
  `_pd_calc_intensity_net`
  `_pd_calc_intensity_total`
  `_pd_calc_point_id`

(*b*) PD_CALC
  `_pd_calc_method`

It is common to calculate powder-diffraction intensities from a crystallographic model. This is necessary for Rietveld refinements, where the model is fitted to the experimentally observed intensities. It is also used to simulate the diffraction pattern of a material for which the structure is known, perhaps for comparison with a measured diffraction pattern.

A crystallographic model can be described in CIF using data items from the core CIF dictionary, as described in Chapter 3.2. To record the results of the simulation, the data items `_pd_calc_intensity_net` or `_pd_calc_intensity_total` are used. The difference between these two data items depends on the treatment of background. If the pattern is simulated with a fitted background added to it, `_pd_calc_intensity_total` is used; otherwise `_pd_calc_intensity_net` is used. The values will typically be placed in a loop with the processed (`_pd_proc_*`) data items or the observed (`_pd_meas_*`) data items. If neither observed

nor processed data are present (*e.g.* for a simulation), or if, for some reason, the simulation has been performed with a different 2θ range or step size, the appropriate `_pd_proc_*` data items are used to define the 2θ values *etc.* used for the simulation.

3.3.5.3. Diffraction peak table

The data items in these categories are as follows:
(*a*) PD_PEAK
● `_pd_peak_id`
  `_pd_peak_2theta_centroid`
  `_pd_peak_2theta_maximum`
  `_pd_peak_d_spacing`
  `_pd_peak_intensity`
  `_pd_peak_pk_height`
  `_pd_peak_wavelength_id`
      → *`_diffrn_radiation_wavelength_id`*
  `_pd_peak_width_2theta`
  `_pd_peak_width_d_spacing`

(*b*) PD_PEAK_METHOD
  `_pd_peak_special_details`

The bullet (●) indicates a category key. The arrow (→) is a reference to a parent data item. Items in italics are defined in the core CIF dictionary.

When diffraction intensities are first measured, particularly when attempting to identify unknown phases in a material, the first step in the analysis is often to compile a list of peak positions. These peak positions are commonly used to search the Powder Diffraction File, which contains lists of peak heights and positions for approximately 100 000 materials (International Centre for Diffraction Data, 2004).

Information on diffraction peaks is recorded in the PD_PEAK section of the pdCIF. Peak positions are recorded using `_pd_peak_2theta_maximum` or `_pd_peak_2theta_centroid`, for positions determined from the intensity maxima or from the peak centroids, respectively. It is also possible to record peak positions using `_pd_peak_d_spacing`. Peak intensities are recorded using `_pd_peak_intensity` and `_pd_peak_pk_height`, for the integrated peak area or the intensity value at the peak maximum, respectively. Peak widths are recorded using `_pd_peak_width_2theta` and `_pd_peak_width_d_spacing`.

A separate loop is used to list reflections, as will be discussed in Section 3.3.5.4. To link reflections to peaks (one peak may consist of many reflections), each peak is assigned a unique code using `_pd_peak_id`, which is then referenced in the reflection table using `_pd_refln_peak_id`.

When intensities are measured using radiation with more than one wavelength, for example when both Cu $K\alpha_1$ and $K\alpha_2$ radiation are used or when a monochromator passes both λ and $\lambda/2$ radiation, peaks may be assigned a wavelength symbol using `_pd_peak_wavelength_id`, where the wavelength symbol is defined in a separate `_diffrn_radiation_wavelength_id` loop. However, for many experiments, the assignment of wavelengths to peaks will be impractical owing to reflection overlap. It is usually better practice to specify wavelength labels in the reflection table using `_pd_refln_wavelength_id`.

3.3.5.4. Reflection assignments and intensities

In addition to the REFLN data items defined in the core CIF dictionary, the following items are defined:
REFLN
  `_pd_refln_peak_id`
      → `_pd_peak_id`
  `_pd_refln_phase_id`
      → `_pd_phase_id`

† `_pd_refln_wavelength_id` (`_refln_wavelength_id`)
          → `_diffrn_radiation_wavelength_id`

The arrow (→) is a reference to a parent data item. The dagger (†) indicates a deprecated item, which should not be used in the creation of new CIFs. Items in italics are defined in the core CIF dictionary.

In a single-crystal experiment, a reflection table contains the initial experimental observations for structural analysis. In contrast, the reflection table for a powder-diffraction experiment is a derived result that depends on the model used to apportion intensity between overlapping reflections. Another difference is that in a single-crystal experiment, the reflection list will refer to only one phase (one hopes), while it is common to have reflections from more than one phase in a powder-diffraction reflection list.

A list of reflections in a powder-diffraction pattern is commonly generated by Rietveld analysis, where Hugo Rietveld's algorithm (Rietveld, 1967, 1969) is used to estimate the intensity of each reflection. Alternatively, when the structure of one or more phases is not known, it is possible to use full-pattern intensity-extraction methods such as the algorithms developed by Pawley (1981) or Le Bail *et al.* (1988). In fact, intensity information obtained by full-pattern intensity extraction is often used for *ab initio* structure determination.

Most of the information in the reflection table will be defined using data items from the core CIF dictionary (see Section 3.2.2.2 and Chapter 4.1). For example, `_refln_index_h`, `_refln_index_k` and `_refln_index_l` will be used for the indices. The structure factors and reflection intensities are specified using `_refln_intensity_calc`, `_refln_intensity_meas`, `_refln_F_squared_calc` and `_refln_F_squared_meas`; reflection positions are defined using `_refln_d_spacing`. To link a reflection with a powder-diffraction peak, the pdCIF data item `_pd_refln_peak_id` is used. The value for `_pd_refln_peak_id` serves as a pointer to an entry in the peak table which has been labelled, using the data name `_pd_peak_id`, with the same symbol. Likewise, to link a reflection to a phase, the pdCIF data item `_pd_refln_phase_id` points to a phase defined using `_pd_phase_id` in the phase table. Since a single reflection may be observed with more than one wavelength, for example, with $\lambda/2$ or $K\alpha_2$ wavelengths, the pdCIF dictionary defines a wavelength link, `_pd_refln_wavelength_id`, that defines a wavelength label. However, since version 2.1, the core CIF dictionary defines `_refln_wavelength_id` and this should be used in preference to `_pd_refln_wavelength_id`. The data items `_refln_wavelength_id` and `_pd_refln_wavelength_id` both point to a wavelength label defined using `_diffrn_radiation_wavelength_id`.

The International Centre for Diffraction Data abstracts peak positions and heights for inclusion in the Powder Diffraction File. This information would be found in the `_pd_peak` section of a pdCIF. However, in many studies, particularly in Rietveld refinements, peak tables are never generated. In principle, it should be possible to calculate peak positions and peak heights (or better still, peak areas) from the information in a reflection table. An algorithm for this would be very useful.

3.3.6. Atomicity, chemistry and structure

The structural model of a compound determined by powder-diffraction methods can be described by the data items in the core CIF dictionary. However, for a powder-diffraction study of a mixture of phases, the PD_PHASE category is used to list the phases present. This is the only category in the pdCIF dictionary that extends the description of the structural model beyond that covered by items in the core CIF dictionary.

3.3.6.1. Table of phases

The data items in this category are as follows:

PD_PHASE
● `_pd_phase_id`
 `_pd_phase_block_id`
 `_pd_phase_mass_%`
 `_pd_phase_name`

The bullet (●) indicates a category key.

When a sample contains more than one phase, the PD_PHASE data items are used to create a table describing the phases present. For example, the name and abundance of each phase can be specified using `_pd_phase_name` and `_pd_phase_mass_%`, respectively.

Two types of pointers can also be defined:

(i) Since the crystallographic description of each phase must be incorporated in a separate data block, `_pd_phase_block_id` contains the unique block ID (see Section 3.3.7) pointing to the block containing the data for the phase.

(ii) An arbitrary label is assigned to every phase using `_pd_phase_id` so that reflections can be assigned to a phase using `_pd_refln_phase_id`. This is discussed further in Section 3.3.5.4.

3.3.7. File metadata

The many data items in the core dictionary that describe file auditing and history cover most of the metadata requirements of a pdCIF, but two new data items in the pdCIF category PD_BLOCK are introduced to provide a specific mechanism for identifying and relating individual data blocks.

Data items in this category are as follows:

PD_BLOCK
 `_pd_block_diffractogram_id`
 `_pd_block_id`

The data item `_pd_block_id` is used to define a unique name for each data block. This name is used so that one data block may reference another data block. Since CIF blocks may be separated into different files, or many CIFs from different sources may be grouped into a single file, the block ID provides a robust mechanism for maintaining references between blocks, independent of how CIF blocks have been arranged between files. The intent is that a site that archives pdCIFs will construct an index to `_pd_block_id` names that can be used to resolve block ID references.

The definition for `_pd_block_id` gives a procedure for creating a `_pd_block_id` name that is extremely unlikely to be duplicated. Other mechanisms for creating unique names can also be used: for example, using a web page name (URL) could be appropriate if care is taken never to reuse the URL.

The need for the block ID/block pointer mechanism is demonstrated by the following example. Consider a case where a neutron powder diffraction data set and an X-ray powder diffraction data set have been used together to determine a single structural model for a single crystalline phase. CIF does not allow the two data sets to be placed in a single block, since this would require two independent loops of observations where each loop uses some of the same data names. One can create a CIF with two blocks and include the structural model in the block that contains either of the two data sets. However, if this is done, a logical link is needed between the two blocks to make it clear that the structural model was derived from both data sets. It is better practice to place the

<table>
<tr><td valign="top">

Example 3.3.7.1. *A CIF with multiple data blocks, demonstrating a suitable construction when multiple data sets and multiple phases occur together.*

```
#= First CIF block ================================
data_NISI_overall

_pd_block_id  2003-02-04T18:02|NISI|B_H_Toby|Overall

# publication and sample preparation information
# appears here (_publ_*, _journal_*, _pd_char_*
# & _pd_prep_* items are omitted for brevity)

# Overall powder R-factors
_pd_proc_ls_prof_wR_factor            0.0370
# (other _refine_ls_* items omitted for brevity)

# pointers to the phase blocks
loop_    _pd_phase_block_id
      2003-02-04T18:02|NISI_phase1|B_H_Toby||
      2003-02-04T18:02|NISI_phase2|B_H_Toby||
# pointers to the diffraction patterns
loop_    _pd_block_diffractogram_id
      2003-02-04T18:02|NISI_H_01|B_H_Toby|GPD
      2003-02-04T18:02|NISI_H_02|B_H_Toby|GPD
```

</td><td valign="top">

Example 3.3.7.1. (*cont.*)

```
#= Second CIF block ================================
# Information for phase 1
data_NISI_phase_1

_pd_block_id 2003-02-04T18:02|NISI_phase1|B_H_Toby||

# Data sets for phase 1
loop_    _pd_block_diffractogram_id
  2003-02-04T18:02|NISI_H_01|B_H_Toby|GPD
  2003-02-04T18:02|NISI_H_02|B_H_Toby|GPD

_pd_phase_name                    Nickel
_cell_length_a                    3.523433(29)
_cell_length_b                    3.523433
_cell_length_c                    3.523433
_cell_angle_alpha                 90.0
_cell_angle_beta                  90.0
_cell_angle_gamma                 90.0
_cell_volume                      43.74194
_symmetry_cell_setting            cubic
_symmetry_space_group_name_H-M    "F m 3 m"
loop_ _symmetry_equiv_pos_site_id
      _symmetry_equiv_pos_as_xyz
      1 +x,+y,+z            2 -x,-y,-z
# (other symmetry operations omitted for brevity)

loop_
      _atom_site_type_symbol
      _atom_site_fract_x
      _atom_site_fract_y
      _atom_site_fract_z
      _atom_site_occupancy
      _atom_site_thermal_displace_type
      _atom_site_U_iso_or_equiv
      _atom_site_symmetry_multiplicity
NI    0.0  0.0  0.0   1.0   Uiso 0.00435(10) 4

loop_  _atom_type_symbol  _atom_type_number_in_cell
      NI  4.0
# (_chemical_* & _geom_* items omitted for brevity)

#= Third CIF block ================================
# Information for phase 2
data_NISI_phase_2

_pd_block_id 2003-02-04T18:02|NISI_phase2|B_H_Toby||

# Data sets for phase 2
loop_    _pd_block_diffractogram_id
  2003-02-04T18:02|NISI_H_01|B_H_Toby|GPD
  2003-02-04T18:02|NISI_H_02|B_H_Toby|GPD

_pd_phase_name                    Silicon
_cell_length_a                    5.42957(9)
_cell_length_b                    5.42957
_cell_length_c                    5.42957
_cell_angle_alpha                 90.0
_cell_angle_beta                  90.0
_cell_angle_gamma                 90.0
_cell_volume                      160.06508
_symmetry_cell_setting            cubic
_symmetry_space_group_name_H-M    "F d 3 m"
loop_ _symmetry_equiv_pos_site_id
      _symmetry_equiv_pos_as_xyz
      1 +x,+y,+z            2 -x,-y,-z
# (other symmetry operations omitted for brevity)

loop_
      _atom_site_type_symbol
      _atom_site_fract_x
      _atom_site_fract_y
      _atom_site_fract_z
      _atom_site_occupancy
      _atom_site_thermal_displace_type
      _atom_site_U_iso_or_equiv
      _atom_site_symmetry_multiplicity
SI 0.125 0.125 0.125  1.0  Uiso  0.00540(21)  8

loop_  _atom_type_symbol  _atom_type_number_in_cell
      SI  8.0
 # (_chemical_* & _geom_* items omitted for brevity)
```

</td></tr>
</table>

structural model in a third data block, as this emphasizes the fact that the model is derived from both data sets. Again, logical links to the data sets are needed.

In both these cases, the data item **_pd_block_diffractogram_id** would be included in the data block containing the structural model and will point to **_pd_block_id** values assigned in the data blocks containing the diffraction data to establish the connection between the data sets and the structural model. The presence of more than one value for **_pd_block_diffractogram_id**, through use of a loop, indicates that multiple data sets were used and thus these structural results are from a combined refinement. Sometimes, powder and single-crystal diffraction data are used together (most commonly to team X-ray single-crystal diffraction data with neutron powder diffraction data). In this case, **_pd_block_diffractogram_id** will point to two **_pd_block_id** values, where one is assigned to the single-crystal data set.

In contrast to the example above, in which block pointers are used to link a single structural model to multiple data sets, another application for these pointers is for describing materials that contain more than one phase. In this case, **_pd_phase_block_id** is placed in the data block containing the data set to link it to the blocks defining the phases.

In summary, three types of links between data blocks are defined.

(i) **_pd_block_diffractogram_id** connects a phase to one or more data-set blocks;

(ii) **_pd_phase_block_id** connects a data set to one or more phase blocks;

(iii) **_pd_calib_std_external_block_id** connects a block to measurements used to provide calibration constants used in the block.

It is good practice to use both **_pd_block_diffractogram_id** and **_pd_phase_block_id** in a pdCIF with multiple blocks.

3.3.7.1. Use of block pointers

More complex link structures will be needed when multiple data sets and multiple phases occur together. Example 3.3.7.1 outlines a pdCIF reporting the results of a TOF powder-diffraction study of a physical mixture of nickel and silicon powders in which two separate diffraction banks, measured at two different Bragg angles, were used. In this case, five CIF blocks are used. The first CIF

<div style="border: 1px solid">

Example 3.3.7.1. (*cont.*)

```
#= Fourth CIF block ================================
# Powder diffraction data for data set 1
data_NISI_p_01

_pd_block_id 2003-02-04T18:02|NISI_H_01|B_H_Toby|GPD

# (numerous _exptl_, _pd_*, _diffrn_ items describing
# the data set are omitted for brevity)

# phase table
loop_  _pd_phase_id
      _pd_phase_block_id
      _pd_phase_mass_%
      _pd_proc_ls_peak_cutoff
  1  2003-02-04T18:02|NISI_phase1|B_H_Toby||
     51(49)    0.00500
  2  2003-02-04T18:02|NISI_phase2|B_H_Toby||
     49(49)    0.00500
#  (_pd_proc_ls_profile_function omitted from loop)

loop_  _atom_type_symbol
      _atom_type_scat_length_neutron
      _atom_type_scat_source
  NI   1.0300   International_Tables_Vol_C
  SI   0.4149   International_Tables_Vol_C

_diffrn_radiation_probe                neutron
_pd_proc_ls_prof_wR_factor             0.0384
_pd_proc_ls_prof_wR_expected           0.0294
_refine_ls_R_Fsqd_factor               0.07288

_pd_proc_info_datetime        2003-02-04T18:02:09
_pd_calc_method               "Rietveld Refinement"
_pd_meas_2theta_fixed         148.29

#---- raw data loop -----
loop_  _pd_meas_time_of_flight
      _pd_meas_intensity_total
      _pd_meas_point_id
      1000.0    1818(34)      626
# (4494 TOF & intensity values omitted for brevity)
_pd_meas_number_of_points            4495

#---- calculated data loop -----
loop_
      _pd_proc_d_spacing
      _pd_proc_intensity_total
      _pd_proc_ls_weight
      _pd_proc_intensity_bkg_calc
      _pd_calc_intensity_total
      _pd_proc_point_id
  0.50035  0.424(7)  19401.  0.3726  0.4155    1
# (1647 processed/calculated points omitted for
# brevity)
_pd_proc_number_of_points            1648

# reflection table
# Note: contains reflections for both phases
loop_
      _refln_index_h
      _refln_index_k
      _refln_index_l
      _pd_refln_phase_id
      _refln_observed_status
      _refln_F_squared_meas
      _refln_F_squared_calc
      _refln_phase_calc
      _refln_d_spacing
  4   0   0  2 o   9.773   9.812  180.00   1.35739
  3   3   1  2 o   4.799   4.801    0.00   1.24563
  2   2   0  1 o  15.254  15.195    0.00   1.24572
# (54 reflections omitted for brevity)
  4   4   4  1 o   7.498   8.733    0.00   0.50856
  9   5   3  2 o   2.350   2.396    0.00   0.50631
  8   6   4  2 o   0.000   0.000  180.00   0.50412
_reflns_number_observed                  60

# (_reflns_limit_* and _reflns_d_* items omitted for
# brevity)
```

</div>

<div style="border: 1px solid">

Example 3.3.7.1. (*cont.*)

```
#= Fifth CIF block ================================
# Powder diffraction data for data set 2
data_NISI_p_02

_pd_block_id 2003-02-04T18:02|NISI_H_02|B_H_Toby|GPD

# (numerous _exptl_, _pd_*, _diffrn_ items describing
# the data set are omitted for brevity)

# phase table
loop_  _pd_phase_id
      _pd_phase_block_id
      _pd_phase_mass_%
      _pd_proc_ls_peak_cutoff
  1  2003-02-04T18:02|NISI_phase1|B_H_Toby||
     51.38       0.00500
  2  2003-02-04T18:02|NISI_phase2|B_H_Toby||
     48.62(28)       0.00500
#  (_pd_proc_ls_profile_function omitted from loop)

loop_  _atom_type_symbol
      _atom_type_scat_length_neutron
      _atom_type_scat_source
  NI   1.0300   International_Tables_Vol_C
  SI   0.4149   International_Tables_Vol_C

_diffrn_radiation_probe                neutron
_pd_proc_ls_prof_wR_factor             0.0363
_pd_proc_ls_prof_wR_expected           0.0222
_refine_ls_R_Fsqd_factor               0.07645

_pd_proc_info_datetime        2003-02-04T18:02:09
_pd_calc_method               "Rietveld Refinement"
_pd_meas_2theta_fixed         88.05

#---- raw data loop -----
loop_  _pd_meas_time_of_flight
      _pd_meas_intensity_total
      _pd_meas_point_id
      750.4    2780(42)      470
# (4650 TOF & intensity values omitted for brevity)
_pd_meas_number_of_points            4651

#---- calculated data loop -----
loop_
      _pd_proc_d_spacing
      _pd_proc_intensity_total
      _pd_proc_ls_weight
      _pd_proc_intensity_bkg_calc
      _pd_calc_intensity_total
      _pd_proc_point_id
  0.45802  0.778(9)  12931.  0.4211  0.7851    1
# (1932 processed/calculated points omitted for
# brevity)
_pd_proc_number_of_points            1933

# reflection table
loop_
      _refln_index_h
      _refln_index_k
      _refln_index_l
      _pd_refln_phase_id
      _refln_observed_status
      _refln_F_squared_meas
      _refln_F_squared_calc
      _refln_phase_calc
      _refln_d_spacing
  2   0   0  1 o  16.505  16.060    0.00   1.76172
  3   1   1  2 o   4.854   5.087  180.00   1.63708
  2   2   2  2 o   0.000   0.000    0.00   1.56738
# (76 reflections omitted for brevity)
  11  3   3  2 o   1.948   2.014    0.00   0.46053
  10  6   2  2 o   0.000   0.000    0.00   0.45888
  7   3   1  1 o   7.261   7.499    0.00   0.45871
  5   5   3  1 o   7.261   7.499    0.00   0.45871
_reflns_number_observed                  83

# (_reflns_limit_* and _reflns_d_* items omitted for
# brevity)
```

</div>

block reports the overall and publication details. The next two CIF blocks report crystallographic information for each phase and the last two blocks report the observed, processed and calculated diffraction intensities and reflection tables.

A second purpose for `_pd_block_id` is to provide a mechanism for tracking successive modifications to a CIF. Consider the case where a data set is obtained at a user facility and the resulting

measurements are distributed as a CIF. In this file, a value is supplied for `_pd_block_id` based on the time when the measurements were made. At a later time, when these observations are analysed, a new CIF is created, containing both the original measurements and the results from the analysis. Rather than replace the original value for `_pd_block_id`, the data item can be placed in a loop and another value, defining a second block ID, can be added. This will

indicate the connection to the initial CIF, since the original block ID is retained.

A potential future use for block pointers may be to reference non-CIF data files that contain large two- and three-dimensional data structures. This is expected to become increasingly important as neutron and synchrotron instruments are constructed that cover increasing ranges of solid angle. As mentioned in Section 3.3.2, CIF is not well suited to these complex, large and possibly irregular measurement arrays. The NeXus format has been developed by a consortium of synchrotron and neutron laboratories to address these concerns and is currently being used for a variety of scattering applications (NeXus, 1999). The NeXus format is based on the platform-independent HDF binary standard (HDF, 1998). The use of block pointers to resolve references to non-CIF documents will require additional definitions.

3.3.8. pdCIF for storing unprocessed measurements

While many researchers prepare a CIF only when a project is complete, there are good reasons for preparing a pdCIF when the diffraction data are measured, as this is the best time to document how the measurement was performed. Much of the instrumental information will remain unchanged for all pdCIFs from a given diffraction instrument, so it is a good idea to prepare a file that describes each of the common settings for an instrument. This file will probably contain some of the following data items and their associated values:

(i) The `_pd_instr_*` items, such as the instrument type in `_pd_instr_geometry`, the size of the instrument and the collimation in `_pd_instr_dist_*` and `_pd_instr_divg_*`, and monochromatization in `_pd_instr_monochr_*` (see Section 3.3.4.3)

(ii) Depending on how the calibration is performed, it may be appropriate to include `_pd_calib_*` items.

(iii) Information about the radiation source should be specified using the `_diffrn_radiation_*` and `_diffrn_source_*` data items.

(iv) Detector information should be specified using `_diffrn_detector_*` items, for example, the detector type in `_diffrn_detector_type` and perhaps calibration values such as the deadtime (in `_diffrn_detector_dtime`).

A second section of the pdCIF will contain information specific to the experiment, such as the diffraction conditions (*i.e.* pressure and temperature) recorded using the `_diffrn_ambient_*` data items. Sample and specimen information will appear in the `_pd_prep_*`, `_pd_spec_*` and `_pd_char_*` data items.

A third section of the pdCIF contains the observations. The data items used to specify the unprocessed observations will vary with the type of instrument used, as described in Sections 3.3.8.1 to 3.3.8.10 below.

3.3.8.1. Single pulse-counting detectors

In the most common measurement method, where a single pulse-counting detector is scanned over a range of 2θ, the `_pd_meas_*` entries (see Section 3.3.4.4) will be of the form shown in Example 3.3.8.1. If the data were scanned using a variable step size, the observations might be given as shown in Example 3.3.8.2. Note that when `_pd_meas_counts_*` is used, the values given must be counts, so that the standard uncertainty will be the square root of the intensity values. This means that the intensity values must not be scaled, for example if the values were counts per second; otherwise the statistical uncertainty estimates will be incorrect.

Example 3.3.8.1. *Measurements from a single pulse-counting detector with constant-step scan.*

```
_pd_meas_2theta_range_min   5.0
_pd_meas_2theta_range_max  65.0
_pd_meas_2theta_range_inc   0.02
_pd_meas_number_of_points  3001
_pd_meas_scan_method       step
_pd_meas_step_count_time   10
loop_
  _pd_meas_counts_total
  10 16 23 18 30 45 58 123 80 67 32 21 12 ...
```

Example 3.3.8.2. *Measurements from a single pulse-counting detector with variable-step scan.*

```
_pd_meas_number_of_points  3001
_pd_meas_scan_method       step
_pd_meas_step_count_time   10
loop_
  _pd_meas_2theta_scan
  _pd_meas_counts_total
  5.00 10 5.02 16 5.04 23 5.06 18 5.07 30 5.08 45
  ... ...
```

3.3.8.2. Detectors that do not count pulses

When the method used to detect intensities does not count individual quanta as they hit the detector, for example, the digitization of intensities recorded on film or on an imaging plate, or even with data recorded using a detector having a built-in deadtime correction, the standard-uncertainty values are not the square root of the intensities. [Note that when the actual deadtime correction is known, it is best to incorporate this scaling into the monitor value (see `_pd_meas_counts_monitor` in Section 3.3.4.4) or else save the uncorrected measurements and create a second set of corrected intensity values as `_pd_proc_intensity_net` (see Section 3.3.5.1).] The `_pd_meas` entries for an experiment using non-pulse-counting detection will look like the examples given in Section 3.3.8.1, except that the data loop will be in the form

```
loop_
  _pd_meas_intensity_total
  10 16 23 18 30 45 58 123 80 67 32 21 12 ...
```

or

```
loop_
  _pd_meas_2theta_scan
  _pd_meas_intensity_total
  5.00 10 5.02 16 5.04 23 5.06 18 5.07 30 5.08 45
  ... ...
```

If standard uncertainties for the intensity values are known, they can be given using the conventional notation

```
loop_
  _pd_meas_2theta_scan
  _pd_meas_intensity_total
  5.00 10(10) 5.02 16(11) 5.04 23(13) 5.06 18(12)
  5.07 30(18) ...
```

Note that when `_pd_meas_intensity_*` is used, it is best to specify `_pd_meas_units_of_intensity` as well.

3.3.8.3. Multiple detectors

At present, CIF does not offer the ability to construct true multi-dimensional data structures. However, many instruments with multiple detectors produce reasonably tractable numbers of data points. For such instruments, it is possible to include an additional data item, `_pd_meas_detector_id`, in the loop with the data to indicate the detector that made the observation.

Example 3.3.8.3. *Identifying intensities from multiple detectors.*

```
loop_
  _pd_meas_2theta_scan
  _pd_meas_intensity_total
  _pd_meas_detector_id
  5.0  10 A   25.0  16 B  45.0 23 C  65.0 18 D
  5.02 16 A   25.02 30 B  ...  ... ...
```

Example 3.3.8.4. *Measurements from an energy-dispersive X-ray diffraction experiment.*

```
_pd_meas_scan_method      disp
_pd_meas_2theta_fixed     6.5
_pd_calib_2theta_offset   0.1071
loop_
  _pd_meas_detector_id
  _pd_meas_counts_total
  _pd_proc_energy_detection
  _pd_proc_recip_len_Q
    0   180  6114.0 .714
    1   166  6141.2 .717
    2    11  6168.4 .721
    3    11  6195.5 .724
```

Example 3.3.8.5. *Measurements from a neutron time-of-flight diffraction experiment.*

```
_pd_meas_scan_method        tof
loop_
  _pd_meas_detector_id
  _pd_meas_time_of_flight
  _pd_proc_d_spacing
  _pd_meas_counts_total
  _pd_proc_intensity_incident
    88  1101.6  0.251658559  11843  4003
    88  1103.2  0.25202477   11934  4001
    88  1104.8  0.252391011  11906  3999
    88  1106.4  0.252757192  11773  3997
    88  1108.0  0.253123432  11707  3995
    ...
    150 1500.0  0.257258087   6559  3185
    ...

loop_
  _pd_calib_detector_id
  _pd_meas_2theta_fixed
    88  88.05
    150 148.29
```

2θ, `_pd_meas_angle_omega` and `_pd_meas_angle_chi` are defined where needed (Example 3.3.8.5).

TOF data are usually reduced to a small number of 'banks' consisting of intensity as a function of d space or Q, where multiple detectors are summed. Data in this form can be recorded using a loop containing `_pd_proc_d_spacing` and `_pd_proc_intensity_net`. A data block is needed for each bank.

3.3.8.6. Digitized film and image plates

To record intensities from digitized X-ray film or from image plates properly requires the storing of two-dimensional data structures, which in some cases can be accommodated through imgCIF (see Chapters 2.3 and 3.7). However, it is possible to record a one-dimensional scan using `_pd_meas_position` and `_pd_meas_intensity_total` (not `_pd_meas_counts_total`!). `_pd_proc_2theta_corrected` values can then be assigned using calibration information, and they can then be included in the same loop, as in Section 3.3.8.4.

3.3.8.7. Direct background measurements

For some diffraction experiments, particularly for the determination of radial distribution functions, measurements are made for background scattering from the diffraction instrument and from the sample container. When this is done, the values can be included in a single loop using `_pd_meas_counts_background`, `_pd_meas_counts_container` and `_pd_meas_counts_total`.

3.3.8.8. Noting sample orientation

For texture measurements, intensity measurements can be made as a function of different sample setting angles. These setting angles can be specified using `_pd_meas_angle_chi`, `_pd_meas_angle_omega` and `_pd_meas_angle_phi`. The change in these values may be specified by including these data items in the loop with the diffraction intensities. In some cases, it may be more convenient to separate measurements with different setting angles into different blocks. In this case, the values for the setting angle(s) that are invariant will be set outside of a loop.

It is common in powder diffraction to reduce preferred orientation and improve crystallite averaging by rocking or rotating the sample. This is indicated by specifying the axis used for rocking, usually φ for capillary specimens or χ for flat-plate specimens, as `_pd_meas_rocking_axis`. The data item

In Example 3.3.8.3, four detectors placed $20°$ apart are referenced with arbitrarily chosen labels A, B, C and D. Note that the detector characteristics will typically be specified in a separate calibration loop containing terms such as `_pd_calib_detector_id`, `_pd_calib_detector_response` and `_pd_calib_2theta_offset`. The labels given for `_pd_calib_detector_id` should match those in `_pd_meas_detector_id`.

3.3.8.4. Energy-dispersive X-ray detection

For energy-dispersive X-ray diffraction, an X-ray detector is placed at a fixed value of 2θ and a diffractogram is measured on a multichannel analyser. The channel number is then calibrated to yield photon energies. From the energy and 2θ angle, a d-spacing or Q value ($Q = 4\pi \sin\theta/\lambda$) is calculated for each diffraction point. Note that energy, d spacing or Q are not the experimental independent variable. Rather, they result from processing, since calibration information is required. The calibration equation should be described in `_pd_calibration_conversion_eqn`.

In Example 3.3.8.4, the nominal 2θ setting is $6.5°$, but the actual position (determined by prior calibration) is $6.6071°$, so the difference is indicated using a `_pd_calib_2theta_offset` value (see Section 3.3.4.3).

3.3.8.5. Neutron time-of-flight detection

Neutron time-of-flight (TOF) detection in theory should be no different from energy-dispersive X-ray detection, but TOF instruments record complex three-dimensional data structures, where diffraction intensities are recorded as a function of time for as many as several hundred detectors. For some instruments, both the position along the detector and the time of flight are recorded, so there may be effectively thousands of detectors. To add even further complexity, the data may be binned in different time steps for detectors at different 2θ values. CIF is likely to be cumbersome for the storage of unprocessed measurements from TOF instruments, owing to the one-dimensional nature of CIF, but it could be useful to translate files from one binary format to another using CIF as a common intermediate. To do this, a single loop is used for all data points, where each detector (or detector section, in the case of a position-sensitive detector) is assigned a detector ID. In a second loop, the detector ID values are defined. In addition to

`_pd_meas_rocking_angle` is used to record the angular range through which the sample is rocked, where 360 indicates that one complete revolution occurs during each counting period. Numbers greater than 360 are possible.

3.3.8.9. Use of an incident-intensity monitor

For radiation sources for which the intensity may vary, such as synchrotron-radiation sources, the intensity of the incident radiation is measured using an incident-intensity monitor. This value may be specified for every data point using `_pd_meas_counts_monitor` (or `_pd_meas_intensity_monitor`) by including this data item in the loop with the diffraction intensities. For some instruments, counting times are set so that the same number of monitor counts are measured for each data point. If this is the case, `_pd_meas_counts_monitor` will be the same for every data point and need not be included in the loop.

3.3.8.10. Recording detector livetime

The detector deadtime is often more a function of the counting electronics than of the intrinsic properties of the detector. In these circumstances, the counting circuit may provide a gating signal that indicates when the electronics are processing an event *versus* when the circuit is idle and waiting for an event to process. From this gating signal, a detector livetime signal can be generated. Livetime is a better way to correct intensities than applying a deadtime correction, because if appreciable numbers of events are processed but are not counted (for example, counts due to fluorescence), the actual deadtime can be quite high, even though the recorded number of counts can be quite low. To use the livetime signal, the count time can be multiplied by the livetime or the livetime can be treated as a monitor (see Section 3.3.8.9). If an incident-intensity monitor and a livetime are both available, the `_pd_meas_intensity_monitor` value can contain the incident intensity times the livetime.

3.3.9. Use of pdCIF for Rietveld refinement results

One of the major aims of the development of the pdCIF definitions was to be able to communicate the results of Rietveld refinements, and this is expected to be the most common use for pdCIF. To aid the development of software that prepares pdCIF output from Rietveld refinements, this section describes the blocks and loops to be found in a pdCIF, noting variations due to the type of Rietveld refinement. Programmers may also wish to look at the *GSAS2CIF* program, which creates CIFs for a wide range of types of diffraction data, and for multiple data sets and phases (Toby *et al.*, 2003).

It is valuable for the CIF to contain the structural model(s), the observed powder-diffraction intensities and the calculated powder-diffraction intensities so that the fit of the model to the observed diffraction pattern can be viewed graphically. It is the present author's firm belief that it is impossible to judge the quality of a Rietveld refinement by *R* factors or any other numerical metric, since these values describe not just how well the structural model fits the measurements, but also how well the background and peak shape are fitted as well. Very poor models can have good *R* factors and χ^2 values if there is a significant amount of non-Bragg scattering that has been well fitted. On the other hand, with high-resolution observations measured to excellent precision, even trivial imperfections in the peak shapes can result in poor agreement factors. There is no substitute for the visual examination of a plot of the observed and calculated patterns, optimally at more than one magnification level. The program *pdCIFplot* (Toby, 2003) plots the observed and calculated powder-diffraction intensities in a pdCIF

and allows the fit to be examined in more detail than can be provided by a figure showing the whole profile at once.

3.3.9.1. A single phase

When a single set of diffraction measurements is used to model a single phase, a pdCIF will usually contain only one block. There will be several important loops present.

One loop will contain atomic parameters, such as coordinates. The unit cell must also be specified.

A second loop will contain the reflection table.

A third loop will contain the observed (or processed) diffraction measurements and the simulated pattern. Other items that should be included in this loop are the least-squares weights (usually σ^{-2}, where σ is the standard uncertainty) so that it is possible to determine the quality of the fit in individual regions. Weight values of zero can also be used to indicate that data points have been excluded from the refinement. Since background fitting is quite important in Rietveld analysis, it is also valuable to include the background values. Thus, this loop should specify:

(i) the ordinate of the Rietveld plot, using one or more of: `_pd_meas_2theta_scan`, `_pd_meas_time_of_flight`, `_pd_proc_2theta_corrected`, `_pd_proc_d_spacing` or `_pd_proc_recip_len_Q`; alternatively the ordinate can be specified using either `_pd_meas_2theta_range_*` or `_pd_proc_2theta_range_*`, where `_*` is `_min`, `_max` and `_inc` outside the loop.

It is recommended that all CIFs describing the results of a Rietveld refinement include either `_pd_proc_d_spacing` or `_pd_proc_recip_len_Q`.

(ii) The observed (or processed) intensity values, using the items `_pd_meas_counts_total`, `_pd_meas_intensity_total`, `_pd_proc_intensity_total` or `_pd_proc_intensity_net`.

(iii) The background, using the item `_pd_proc_intensity_bkg_calc`.

(iv) The least-squares weights, using the item `_pd_proc_ls_weight`. If these weights are not specified, then it must be presumed that all points have been used in the refinement and that the weights are the reciprocal of the intensity values (if `_pd_meas_counts_total` was used) or the reciprocal of the intensity standard uncertainties, if specified.

(v) The calculated pattern should appear using either `_pd_calc_intensity_net` or `_pd_calc_intensity_total`.

It is good practice always to include at least one data item from each entry in the list above.

Apart from the information contained in these loops, information from almost all sections of the pdCIF dictionary can be valuable. Such items include data items that define how the diffraction measurements were made, how the sample was prepared and characterized, how the refinement was performed, and least-squares parameters and *R* factors. A template and an example pdCIF showing the combined use of pdCIF and core data items form part of the *Acta Crystallographica* instructions for authors at http://journals.iucr.org/services/cif/powder.html.

3.3.9.2. Multiple phases

When more than one phase is present, multiple CIF blocks are needed. The resulting CIF will contain much the same information as would be found in a single-phase pdCIF, as described in Section 3.3.9.1. However, there will be a separate block for each phase containing information specific to that phase, such as the unit cell and the loop containing the atomic parameters.

The CIF will usually (see Section 3.3.7) contain one additional block with the observed and calculated pattern and a reflection table, as well as the other data items that define how the

diffraction measurements were made, how the refinement was performed *etc.* While reflection tables for each phase can be placed in each phase block, it is better to include a single reflection table in the block that contains the diffraction data. This block will also contain a phase table that uses the block pointer `_pd_block_diffractogram_id` to link to the phase blocks. The phase blocks can also be linked to the data block using the block pointer `_pd_phase_block_id`. For most Rietveld refinements, each phase is allowed to have different profile parameters, so `_pd_proc_ls_profile_function` should also be included in the phase-table loop.

3.3.9.3. One phase, multiple sets of measurements

It is fairly common to use more than one diffraction data set to determine a model for a single phase. Some examples include: combined refinement using both neutron and X-ray powder diffraction; use of multiple X-ray wavelengths to make use of anomalous dispersion; and the use of single-crystal X-ray and powder neutron diffraction data in a single refinement. For these cases, there will be a CIF block for each data set. Each of these blocks will contain a reflection table and a loop with the observed and calculated diffraction intensities, as described in Section 3.3.9.1.

As explained in Section 3.3.7, the resulting structural parameters could be placed in a block with one of the sets of diffraction data. However, it is better practice to create one additional block for these parameters, as it then becomes clear that the result is from a combined refinement. This is indicated by linking the phase block and the data-set blocks using a loop of `_pd_block_diffractogram_id` values in the phase block. The data-set blocks can also have a link to the phase information using the block pointer `_pd_phase_block_id`.

3.3.9.4. Multiple sets of measurements and phases

Multiple data sets may be used for mixtures as well as single phases. This is becoming increasingly common as more complex materials are studied using powder diffraction. The treatment of this case follows logically from that of Sections 3.3.9.2 and 3.3.9.3. If there are M diffraction data sets and P phases, there will be P blocks containing the crystallographic parameters for each phase. There will be M blocks with the observed and calculated diffraction intensities, as well as reflection tables. Depending on the Rietveld software, there may be $M \times P$ sets of some parameters, for example phase fractions and profile descriptions. These parameters may be placed in the phase table loop within the data-set block(s).

Ideally, the same specimen, or at least the same sample, will be used for all measurements. Sometimes, however, different samples are used for combined refinements to extend the number of observations, despite the possibility that the samples might have slightly different structures or compositions. If there are S samples, there will be an additional S blocks that record the sample and specimen preparation and characterization information. Thus, in this case there will be a total of $M + P + S$ blocks.

As before, the phase blocks will use the block pointers `_pd_block_diffractogram_id` to link to the data-set blocks. Likewise, the data-set blocks will have phase tables with `_pd_phase_block_id` values that link to the phase blocks. The sample blocks can use both `_pd_block_diffractogram_id` values and `_pd_phase_block_id` values to link to the the diffraction data and the analysis results. This is shown in the CIF in Example

---

Example 3.3.10.1. *Phase identification using a reflection table and a phase table.*

```
loop_
  _pd_peak_2theta_centroid
  _pd_peak_id
     3 A1
     4 B1
     6 A2
     8 B2
     9 A3
    12 A4
    15 A5
    16 B3

loop_
   _refln_index_h
   _refln_index_k
   _refln_index_l
   _pd_refln_peak_id
   _pd_refln_phase_id
   ? ? ? A1 Phase1
   ? ? ? A2 Phase1
   ? ? ? A3 Phase1
   ? ? ? A4 Phase1
   ? ? ? A5 Phase1

loop_
   _pd_phase_id
   _pd_phase_name
   Phase1  'component 1'
```

3.3.7.1. The program *GSAS2CIF* (Toby *et al.*, 2003) can create CIFs for multiple sets of measurements and phases.

3.3.10. Other pdCIF applications

As mentioned above, there are other applications for pdCIF than the storage of unprocessed measurements and the reporting of the results of a Rietveld refinement. This section describes the use of data items in other common pdCIF applications.

3.3.10.1. Simulated intensities

It is common to simulate a diffraction pattern from a known or hypothetical structural model. The structural model is recorded in CIF using core data items, such as `_atom_site_label`, `_atom_site_fract_x`, `_atom_site_fract_y`, `_atom_site_fract_z`, `_atom_site_U_iso_or_equiv` and `_atom_site_occupancy`, as well as the unit cell in `_cell_length_*` and `_cell_angle_*`. Calculated reflection intensities can be recorded using `_refln_index_*` and `_refln_F_squared_calc`, as described in Section 3.3.5.4. The simulated pattern can be recorded using `_pd_calc_*` data items, as described in Section 3.3.5.2.

The simulated diffraction pattern will be determined not only by the structural parameters, but also by the type of experiment that is being simulated. For example, it is good practice to define data items to specify the type of radiation in `_diffrn_radiation_probe`, the wavelength in `_diffrn_radiation_wavelength` and the profile in `_pd_proc_ls_profile_function`.

3.3.10.2. Phase identification and indexing

For phase identification, a CIF will include unprocessed measurements, as described in Section 3.3.8. Sample characterization information, for example chemical analysis information, can often aid phase determination. Characterization information is described in Section 3.3.4.1. Similarly, sample preparation information can also be quite valuable (see Section 3.3.4.1). Since preferred orientation or other artifacts of the measurement can make phase

Table A3.3.1.1. *Categories in the powder CIF dictionary*

Numbers in parentheses refer to the section of this chapter where each category is described in detail.

PD_BLOCK group (§3.3.7)	PD_MEAS group (§3.3.4.4)
PD_BLOCK (§3.3.7)	PD_DATA (*part of category*) (§3.3.4.4(*a*))
PD_CALC group (§3.3.5.2)	PD_MEAS_INFO (§3.3.4.4(*b*))
PD_CALC (§3.3.5.2(*b*))	PD_MEAS_METHOD (§3.3.4.4(*c*))
PD_DATA (*part of category*) (§3.3.5.2(*a*))	PD_PEAK group (§3.3.5.3)
PD_CALIB group (§3.3.4.3)	PD_PEAK (§3.3.5.3(*a*))
PD_CALIB (§3.3.4.3(*a*))	PD_PEAK_METHOD (§3.3.5.3(*b*))
PD_CALIBRATION (§3.3.4.3(*b*))	PD_PHASE group (§3.3.6.1)
PD_CHAR group (§3.3.4.1)	PD_PHASE (§3.3.6.1)
PD_CHAR (§3.3.4.1(*a*))	PD_PREP group (§3.3.4.1)
PD_DATA group	PD_PREP (§3.3.4.1(*b*))
`_pd_calc_*` items (§3.3.5.2(*a*))	PD_PROC group (§3.3.5.1)
`_pd_instr_var_illum_len`	PD_DATA (*part of category*) (§3.3.5.1(*a*))
(§3.3.4.3(*d*))	PD_PROC_INFO (§3.3.5.1(*b*))
`_pd_meas_*` items (§3.3.4.4(*a*))	PD_PROC_LS (§3.3.5.1(*c*))
`_pd_proc_*` items (§3.3.5.1(*a*))	PD_SPEC group (§3.3.4.2)
PD_INSTR group (§3.3.4.3)	PD_SPEC (§3.3.4.2)
PD_DATA (*part of category*) (§3.3.4.3(*d*))	REFLN group (§3.3.5.4)
PD_INSTR (§3.3.4.3(*c*))	REFLN (§3.3.5.4)

identification more difficult, it is a good idea to include specimen preparation and mounting information, as described in Section 3.3.4.2.

Peaks will be located and documented in the peak table, as discussed in Section 3.3.5.3. In this case, it can be very helpful to specify `_pd_peak_wavelength_id` for peaks that are clearly a $K\alpha_2$ component. Similarly, recording a peak width can also be helpful for some autoindexing programs.

To identify peaks by phase in the case where at least one phase in a material is known but other peaks remain unidentified requires the use of a reflection table and a phase table, as shown in Example 3.3.10.1. This example shows a diffraction pattern with eight peaks, of which five have been identified as arising from a phase where the unit cell has not yet been determined. The peaks labelled B1, B2 and B3 are referenced in the peak table, but are not defined in the reflection loop. This implies that they arise from an unknown phase or phases. The remaining peaks A1 to A5 are referenced in the peak table. Note that the reflection table must include the `_refln_index_*` data items, even though no reflection indices are assigned. This is because CIF rules require that the `_refln_index_*` data items be present in this loop, as noted in the pdCIF dictionary definitions for `_pd_refln_peak_id` and `_pd_refln_phase_id`. The place holder ? is used to indicate that the reflection indices are not yet known.

Appendix 3.3.1
Category structure of the powder CIF dictionary

Table A3.3.1.1 provides an overview of the structure of the powder CIF dictionary by informal category group.

Many people contributed to the creation of this dictionary, most notably Syd Hall, I. David Brown and Brian McMahon. Ian Langford helped considerably with the evolution of the dictionary. This project also benefited from the efforts of many members of the PDF-3 Database Format Subcommittee of the International Centre for Diffraction Data; I am particularly indebted to Walter Schriener for useful concepts and discussions. My initial involvement with powder diffraction file formats was prompted by the encouragement of Richard Harlow; I may someday forgive him.

References

Finger, L. W., Cox, D. E. & Jephcoat, A. P. (1994). *A correction for powder diffraction peak asymmetry due to axial divergence. J. Appl. Cryst.* **27**, 892–900.

HDF (1998). *NCSA HDF home page*. http://www.hdfgroup.org/.

International Centre for Diffraction Data (2004). Powder Diffraction File. International Centre for Diffraction Data, 12 Campus Boulevard, Newtown Square, Pennsylvania, USA. http://www.icdd.com.

Le Bail, A., Duroy, H. & Fourquet, J. L. (1988). *Ab initio structure determination of LiSbWO$_6$ by X-ray powder diffraction. Mater. Res. Bull.* **23**, 447–452.

NeXus (1999). *NeXus data format*. http://www.nexusformat.org.

Pawley, G. S. (1981). *Unit-cell refinement from powder diffraction scans. J. Appl. Cryst.* **14**, 357–361.

Rietveld, H. M. (1967). *Line profiles of neutron powder-diffraction peaks for structure refinement. Acta Cryst.* **22**, 151–152.

Rietveld, H. M. (1969). *A profile refinement method for nuclear and magnetic structures. J. Appl. Cryst.* **2**, 65–71.

Toby, B. H. (2003). *CIF applications. XII. Inspecting Rietveld fits from pdCIF: pdCIFplot. J. Appl. Cryst.* **36**, 1285–1287.

Toby, B. H., Von Dreele, R. B. & Larson, A. C. (2003). *CIF applications. XIV. Reporting of Rietveld results using pdCIF: GSAS2CIF. J. Appl. Cryst.* **36**, 1290–1294.

references