International Tables for Crystallography (2006). Vol. G, Section 3.3.8, pp. 126–128.

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
```

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

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
```

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 _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

references