International Tables for Crystallography (2006). Vol. G, Section 3.3.7.1, pp. 124–126.

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
```

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

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)
```

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)
```

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)
```

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.

references