Example 3.6.6.5. *Phasing of the structure of bovine plasma retinol-binding protein (Zanotti et al., 1993) described using data items in the* PHASING_MIR *and related categories.*

```
_phasing_MIR.entry_id          '1HBP'
_phasing_MIR.method
;  Standard phase refinement (Blow & Crick, 1959)
;

loop_
_phasing_MIR_shell.d_res_low
_phasing_MIR_shell.d_res_high
_phasing_MIR_shell.reflns
_phasing_MIR_shell.FOM
 15.0  8.3   80  0.69       8.3  6.4  184  0.73
  6.4  5.2  288  0.72       5.2  4.4  406  0.65
  4.4  3.8  554  0.54       3.8  3.4  730  0.53
  3.4  3.0  939  0.50

loop_
_phasing_MIR_der.id
_phasing_MIR_der.number_of_sites
_phasing_MIR_der.details
KAu(CN)2  3
     'major site interpreted in difference Patterson'
K2HgI4    6 'sites found in cross-difference Fourier'
K3IrCl6   2 'sites found in cross-difference Fourier'
All      11 'data for all three derivatives combined'

loop_
_phasing_MIR_der_shell.der_id
_phasing_MIR_der_shell.d_res_low
_phasing_MIR_der_shell.d_res_high
_phasing_MIR_der_shell.ha_ampl
_phasing_MIR_der_shell.loc
   KAu(CN)2  15.0  8.3   54   26
   KAu(CN)2   8.3  6.4   54   20
# - - - abbreviated - - -
   K2HgI4    15.0  8.3  149   87
   K2HgI4     8.3  6.4  121   73
# - - - abbreviated - - -
   K3IrCl6   15.0  8.3   33   27
   K3IrCl6    8.3  6.4   40   23
# - - - abbreviated - - -

loop_
_phasing_MIR_der_site.der_id
_phasing_MIR_der_site.id
_phasing_MIR_der_site.atom_type_symbol
_phasing_MIR_der_site.occupancy
_phasing_MIR_der_site.fract_x
_phasing_MIR_der_site.fract_y
_phasing_MIR_der_site.fract_z
_phasing_MIR_der_site.B_iso
   KAu(CN)2  1  Au  0.40  0.082  0.266  0.615  33.0
   KAu(CN)2  2  Au  0.03  0.607  0.217  0.816  25.9
   K2HgI4    1  Hg  0.63  0.048  0.286  0.636  33.7
   K2HgI4    2  Hg  0.34  0.913  0.768  0.889  36.7
# - - - abbreviated - - -

_phasing_MIR_der_refln.index_h            6
_phasing_MIR_der_refln.index_k            1
_phasing_MIR_der_refln.index_l           25
_phasing_MIR_der_refln.der_id         HGPT1
_phasing_MIR_der_refln.set_id       'NS1-96'
_phasing_MIR_der_refln.F_calc_au     106.66
_phasing_MIR_der_refln.F_meas_au     204.67
_phasing_MIR_der_refln.F_meas_sigma    6.21
_phasing_MIR_der_refln.HL_A_iso       -3.15
_phasing_MIR_der_refln.HL_B_iso       -0.76
_phasing_MIR_der_refln.HL_C_iso        0.65
_phasing_MIR_der_refln.HL_D_iso        0.23
_phasing_MIR_der_refln.phase_calc    194.48
```

3.6.6.1.6. *Phasing data sets*

The data items in these categories are as follows:

(*a*) PHASING_SET

- **_phasing_set.id**

  _phasing_set.cell_angle_alpha
  _phasing_set.cell_angle_beta
  _phasing_set.cell_angle_gamma
  _phasing_set.cell_length_a
  _phasing_set.cell_length_b
  _phasing_set.cell_length_c
  _phasing_set.detector_specific
  _phasing_set.detector_type
  _phasing_set.radiation_source_specific
  _phasing_set.radiation_wavelength
  _phasing_set.temp

(*b*) PHASING_SET_REFLN

- **_phasing_set_refln.index_h**
- **_phasing_set_refln.index_k**
- **_phasing_set_refln.index_l**
- **_phasing_set_refln.set_id**
  → _phasing_set.id

  _phasing_set_refln.F_meas
  _phasing_set_refln.F_meas_au
  _phasing_set_refln.F_meas_sigma
  _phasing_set_refln.F_meas_sigma_au

The bullet (•) *indicates a category key. Where multiple items within a category are marked with a bullet, they must be taken together to form a compound key. The arrow* (→) *is a reference to a parent data item.*

Data items in the PHASING_SET family of categories are homologous to items with related names in the CELL and DIFFRN families of categories. The PHASING_SET categories were added to the mmCIF data model so that intensity and phase information for the data sets used in phasing could be stored in the same data block as the information for the refined structure. It is not necessary to store all the experimental information for each data set (*e.g.* the raw data sets or crystal growth conditions); it is assumed that the full experimental description of each phasing set would be recorded in a separate data block (see Example 3.6.6.6).

Data items in the PHASING_SET category identify each set of diffraction data used in a phasing experiment and can be used to summarize relevant experimental conditions. Because a given data set may be used in a number of different ways (for example, as an isomorphous derivative and as a component of a multiple-wavelength calculation), it is appropriate to store the reflections in a category distinct from either the PHASING_MAD or PHASING_MIR family of categories, but accessible to both these families (and any similar categories that might be introduced later to describe new phasing methods). Figs. 3.6.6.1 and 3.6.6.2 show how reference is made to the relevant sets from within the PHASING_MAD and PHASING_MIR categories.

Each phasing set is given a unique value of **_phasing_set.id**. The other PHASING_SET data items record the cell dimensions and

Example 3.6.6.6. *The phasing sets used in the structure determination of bovine plasma retinol-binding protein (Zanotti et al., 1993) described with data items in the* PHASING_SET *and* PHASING_SET_REFLN *categories.*

```
_phasing_set.id                    'NS1-96'
_phasing_set.cell_angle_alpha      90.0
_phasing_set.cell_angle_beta       90.0
_phasing_set.cell_angle_gamma      90.0
_phasing_set.cell_length_a         38.63
_phasing_set.cell_length_b         38.63
_phasing_set.cell_length_c         82.88
_phasing_set.radiation_wavelength  1.5145
_phasing_set.detector_type         'image plate'
_phasing_set.detector_specific     'RXII'

_loop
_phasing_set_refln.set_id
_phasing_set_refln.index_h
_phasing_set_refln.index_k
_phasing_set_refln.index_l
_phasing_set_refln.F_meas_au
_phasing_set_refln.F_meas_sigma_au
  'NS1-96'  15  15  32  181.79  3.72
  'NS1-96'  15  15  33   34.23  1.62
# - - - abbreviated - - -
```

angles associated with each phasing set, the wavelength of the radiation used in the experiment, the source of the radiation, the detector type, and the ambient temperature.

Data items in the PHASING_SET_REFLN category are used to record the values of the measured structure factors and their uncertainties. Several distinct data sets may be present in this list, with reflections in each set identified by the appropriate value of **_phasing_set_refln.set_id**.

3.6.6.2. Refinement

The categories describing refinement are as follows:

REFINE group

Overall description of the refinement (§3.6.6.2.1)

REFINE

REFINE_FUNCT_MINIMIZED

Analysis of the refined structure (§3.6.6.2.2)

REFINE_ANALYZE

Restraints and refinement by shells of resolution (§3.6.6.2.3)

REFINE_LS_RESTR

REFINE_LS_RESTR_NCS

REFINE_LS_RESTR_TYPE

REFINE_LS_SHELL

REFINE_LS_CLASS

Equivalent atoms in the refinement (§3.6.6.2.4)

REFINE_B_ISO

REFINE_OCCUPANCY

History of the refinement (§3.6.6.2.5)

REFINE_HIST

The macromolecular CIF dictionary contains many more data items for describing the refinement process than the core CIF dictionary does. In addition to new items in the REFINE category itself, additional categories have been introduced to describe in great detail the function minimized and the restraints applied, and the history of the refinement process, which often has many cycles. The REFINE_ANALYZE category can be used to give details of many of the quantities that may be used to assess the quality of the refinement. The REFINE_LS_SHELL category allows results to be reported by shells of resolution, and in effect replaces the more general core CIF category REFINE_LS_CLASS.

3.6.6.2.1. *Overall description of the refinement*

The data items in these categories are as follows:

(*a*) REFINE

- **_refine.entry_id**
  → **_entry.id**
  _refine.aniso_B[1][1]
  _refine.aniso_B[1][2]
  _refine.aniso_B[1][3]
  _refine.aniso_B[2][2]
  _refine.aniso_B[2][3]
  _refine.aniso_B[3][3]
  _refine.B_iso_max
  _refine.B_iso_mean
  _refine.B_iso_min
  _refine.correlation_coeff_Fo_to_Fc
  _refine.correlation_coeff_Fo_to_Fc_free
  _refine.details (∼ ***_refine_special_details***)
+ ***_refine.diff_density_max***
+ ***_refine.diff_density_min***
+ ***_refine.diff_density_rms***
  _refine.ls_abs_structure_details
+ ***_refine.ls_abs_structure_Flack***
+ ***_refine.ls_abs_structure_Rogers***
  _refine.ls_d_res_high
  _refine.ls_d_res_low
+ ***_refine.ls_extinction_coef***
  _refine.ls_extinction_expression
  _refine.ls_extinction_method
+ ***_refine.ls_goodness_of_fit_all***
  _refine.ls_goodness_of_fit_gt
+ ***_refine.ls_goodness_of_fit_obs***
  _refine.ls_goodness_of_fit_ref
  _refine.ls_hydrogen_treatment
  _refine.ls_matrix_type
  _refine.ls_number_constraints
  _refine.ls_number_parameters
  _refine.ls_number_reflns_all
  _refine.ls_number_reflns_obs
  (∼ ***_refine_ls_number_reflns***)
  _refine.ls_number_reflns_R_free
  _refine.ls_number_reflns_R_work
  _refine.ls_number_restraints
  _refine.ls_percent_reflns_obs
  _refine.ls_percent_reflns_R_free
  _refine.ls_R_factor_all
  _refine.ls_R_factor_gt
  _refine.ls_R_factor_obs
  _refine.ls_R_factor_R_free
  _refine.ls_R_factor_R_free_error
  _refine.ls_R_factor_R_free_error_details
  _refine.ls_R_factor_R_work
  _refine.ls_R_Fsqd_factor_obs
  (∼ ***_refine_ls_R_Fsqd_factor***)
  _refine.ls_R_I_factor_obs (∼ ***_refine_ls_R_I_factor***)
  _refine.ls_redundancy_reflns_all
  _refine.ls_redundancy_reflns_obs
  _refine.ls_restrained_S_all
  _refine.ls_restrained_S_obs
  _refine.ls_shift_over_esd_max
  (∼ ***_refine_ls_shift/esd_max***)
  _refine.ls_shift_over_esd_mean
  (∼ ***_refine_ls_shift/esd_mean***)
  _refine.ls_shift_over_su_max
  (∼ ***_refine_ls_shift/su_max***)
  _refine.ls_shift_over_su_max_lt
  (∼ ***_refine_ls_shift/su_max_lt***)
  _refine.ls_shift_over_su_mean
  (∼ ***_refine_ls_shift/su_mean***)
  _refine.ls_shift_over_su_mean_lt
  (∼ ***_refine_ls_shift/su_mean_lt***)
  _refine.ls_structure_factor_coef
  _refine.ls_weighting_details
  _refine.ls_weighting_scheme
  _refine.ls_wR_factor_all
  _refine.ls_wR_factor_obs
  _refine.ls_wR_factor_R_free
  _refine.ls_wR_factor_R_work
  _refine.occupancy_max
  _refine.occupancy_min
  _refine.overall_FOM_free_R_set
  _refine.overall_FOM_work_R_set
  _refine.overall_SU_B
  _refine.overall_SU_ML
  _refine.overall_SU_R_Cruickshank_DPI
  _refine.overall_SU_R_free
  _refine.solvent_model_details
  _refine.solvent_model_param_bsol
  _refine.solvent_model_param_ksol

(*b*) REFINE_FUNCT_MINIMIZED

- **_refine_funct_minimized.type**
  _refine_funct_minimized.number_terms
  _refine_funct_minimized.residual
  _refine_funct_minimized.weight

The bullet (•) indicates a category key. The arrow (→) is a reference to a parent data item. Items in italics have aliases in the core CIF dictionary formed by changing the full stop (.) to an underscore (_) except where indicated by the ∼ symbol. Data items marked with a plus (+) have companion data names for the standard uncertainty in the reported value, formed by appending the string **_esd** *to the data name listed.*

There is already an extensive set of data names in the REFINE category of the core dictionary, and Section 3.2.3.1 should be read with the present section. The only data items discussed in this section are entries in the mmCIF dictionary that do not have a counterpart in the core CIF dictionary. Analogues of a number of *R* factors in the core CIF dictionary have been added to the mmCIF dictionary to express these same *R* factors indepen-