

## 4.5. Macromolecular dictionary (mmCIF)

BY P. M. D. FITZGERALD, J. D. WESTBROOK, P. E. BOURNE, B. MCMAHON, K. D. WATENPAUGH AND H. M. BERMAN

This is version 2.0.09 of the macromolecular CIF dictionary (mmCIF). The philosophy behind this dictionary and the history of its development are described in Chapter 1.1. A detailed commentary on the use of the dictionary is given in Chapter 3.6.

### Category groups

atom_group	Categories that describe the properties of atoms.
audit_group	Categories that describe dictionary maintenance and identification.
cell_group	Categories that describe the unit cell.
chemical_group	Categories that describe chemical properties and nomenclature.
chem_comp_group	Categories that describe components of chemical structure.
chem_link_group	Categories that describe links between components of chemical structure.
citation_group	Categories that provide bibliographic references.
computing_group	Categories that describe the computational details of the experiment.
compliance_group	Categories that are included in this dictionary specifically to comply with previous dictionaries.
database_group	Categories that hold references to entries in databases that contain related information.
diffn_group	Categories that describe details of the diffraction experiment.
entity_group	Categories that describe chemical entities.
entry_group	Categories that pertain to the entire data block.
exptl_group	Categories that hold details of the experimental conditions.
geom_group	Categories that hold details of molecular and crystal geometry.
iucr_group	Categories that are used for manuscript submission and internal processing by the staff of the International Union of Crystallography.
pdb_group	Categories that pertain to the file-format or data-processing codes used by the Protein Data Bank.

Affiliations: PAULA M. D. FITZGERALD, Merck Research Laboratories, Rahway, New Jersey, USA; JOHN D. WESTBROOK, Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers, The State University of New Jersey, Department of Chemistry and Chemical Biology, 610 Taylor Road, Piscataway, NJ, USA; PHILLIP E. BOURNE, Research Collaboratory for Structural Bioinformatics, San Diego Supercomputer Center, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0537, USA; BRIAN MCMAHON, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England; KEITH D. WATENPAUGH, retired; formerly Structural, Analytical and Medicinal Chemistry, Pharmacia Corporation, Kalamazoo, Michigan, USA; HELEN M. BERMAN, Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers, The State University of New Jersey, Department of Chemistry and Chemical Biology, 610 Taylor Road, Piscataway, NJ, USA.

phasing_group	Categories that describe phasing.
refine_group	Categories that describe refinement.
refln_group	Categories that describe the details of reflection measurements.
struct_group	Categories that contain details about the crystallographic structure.
symmetry_group	Categories that describe symmetry information.

### ATOM\_SITE

Data items in the ATOM\_SITE category record details about the atom sites in a macromolecular crystal structure, such as the positional coordinates, atomic displacement parameters, magnetic moments and directions. The data items for describing anisotropic atomic displacement factors are only used if the corresponding items are not given in the ATOM\_SITE\_ANISOTROP category.

Category group(s): **inclusive\_group**  
**atom\_group**

Category key(s): **\_atom\_site.id**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```

loop_
  _atom_site.group_PDB
  _atom_site.type_symbol
  _atom_site.label_atom_id
  _atom_site.label_comp_id
  _atom_site.label_asym_id
  _atom_site.label_seq_id
  _atom_site.label_alt_id
  _atom_site.Cartn_x
  _atom_site.Cartn_y
  _atom_site.Cartn_z
  _atom_site.occupancy
  _atom_site.B_iso_or_equiv
  _atom_site.footnote_id
  _atom_site.auth_seq_id
  _atom_site.id
ATOM N N VAL A 11 . 25.369 30.691 11.795 1.00
17.93 . 11 1
ATOM C CA VAL A 11 . 25.970 31.965 12.332 1.00
17.75 . 11 2
ATOM C C VAL A 11 . 25.569 32.010 13.808 1.00
17.83 . 11 3
ATOM O O VAL A 11 . 24.735 31.190 14.167 1.00
17.53 . 11 4
ATOM C CB VAL A 11 . 25.379 33.146 11.540 1.00
17.66 . 11 5
ATOM C CG1 VAL A 11 . 25.584 33.034 10.030 1.00
18.86 . 11 6
ATOM C CG2 VAL A 11 . 23.933 33.309 11.872 1.00
17.12 . 11 7
ATOM N N THR A 12 . 26.095 32.930 14.590 1.00
18.97 4 12 8
ATOM C CA THR A 12 . 25.734 32.995 16.032 1.00
19.80 4 12 9
ATOM C C THR A 12 . 24.695 34.106 16.113 1.00
20.92 4 12 10
ATOM O O THR A 12 . 24.869 35.118 15.421 1.00
21.84 4 12 11
ATOM C CB THR A 12 . 26.911 33.346 17.018 1.00
20.51 4 12 12
ATOM O OG1 THR A 12 3 27.946 33.921 16.183 0.50
20.29 4 12 13

```

ATOM	O	OG1	THR	A	12	4	27.769	32.142	17.103	0.50
		20.59	4	12	14					
ATOM	C	CG2	THR	A	12	3	27.418	32.181	17.878	0.50
		20.47	4	12	15					
ATOM	C	CG2	THR	A	12	4	26.489	33.778	18.426	0.50
		20.00	4	12	16					
ATOM	N	N	ILE	A	13	.	23.664	33.855	16.884	1.00
		22.08	.	13	17					
ATOM	C	CA	ILE	A	13	.	22.623	34.850	17.093	1.00
		23.44	.	13	18					
ATOM	C	C	ILE	A	13	.	22.657	35.113	18.610	1.00
		25.77	.	13	19					
ATOM	O	O	ILE	A	13	.	23.123	34.250	19.406	1.00
		26.28	.	13	20					
ATOM	C	CB	ILE	A	13	.	21.236	34.463	16.492	1.00
		22.67	.	13	21					
ATOM	C	CG1	ILE	A	13	.	20.478	33.469	17.371	1.00
		22.14	.	13	22					
ATOM	C	CG2	ILE	A	13	.	21.357	33.986	15.016	1.00
		21.75	.	13	23					
# - - - data truncated for brevity - - -										
HETATM	C	C1	APS	C	.	1	4.171	29.012	7.116	0.58
		17.27	1	300	101					
HETATM	C	C2	APS	C	.	1	4.949	27.758	6.793	0.58
		16.95	1	300	102					
HETATM	O	O3	APS	C	.	1	4.800	26.678	7.393	0.58
		16.85	1	300	103					
HETATM	N	N4	APS	C	.	1	5.930	27.841	5.869	0.58
		16.43	1	300	104					
# - - - data truncated for brevity - - -										

atom\_site.adp\_type (code)

atom\_site\_adp\_type (cif\_core.dic 2.3)

A standard code used to describe the type of atomic displacement parameters used for the site.

Related item: atom\_site.thermal\_displace\_type (alternate).

The data value must be one of the following:

Uani anisotropic  $U^{ij}$   
 Uiso isotropic  $U$   
 Uovl overall  $U$   
 Umpe multipole expansion  $U$   
 Bani anisotropic  $B^{ij}$   
 Biso isotropic  $B$   
 Bovl overall  $B$

[atom\_site]

atom\_site.aniso\_B[1][1] (float, su)

The [1][1] element of the anisotropic atomic displacement matrix **B**, which appears in the structure-factor term as

$$T = \exp\left\{-\frac{1}{4} \sum_i \left[ \sum_j (B^{ij} h_i h_j a_i^* a_j^*) \right] \right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row. The IUCr Commission on Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: atom\_site.aniso\_B[1][1]\_esd (associated esd),

atom\_site.aniso\_U[1][1] (conversion constant),  
atom\_site\_anisotrop.U[1][1] (conversion constant),  
atom\_site.aniso\_U[1][1]\_esd (alternate exclusive),  
atom\_site\_anisotrop.B[1][1] (alternate exclusive),  
atom\_site\_anisotrop.U[1][1]\_esd (alternate exclusive).

[atom\_site]

atom\_site.aniso\_B[1][1]\_esd (float)

The standard uncertainty (estimated standard deviation) of atom\_site.aniso\_B[1][1].

Related items: atom\_site.aniso\_B[1][1] (associated value),  
atom\_site.aniso\_U[1][1]\_esd (conversion constant),  
atom\_site\_anisotrop.U[1][1]\_esd (conversion constant),  
atom\_site.aniso\_U[1][1]\_esd (alternate exclusive),  
atom\_site\_anisotrop.B[1][1]\_esd (alternate exclusive),  
atom\_site\_anisotrop.U[1][1]\_esd (alternate exclusive).

[atom\_site]

atom\_site.aniso\_B[1][2] (float, su)

The [1][2] element of the anisotropic atomic displacement matrix **B**, which appears in the structure-factor term as

$$T = \exp\left\{-\frac{1}{4} \sum_i \left[ \sum_j (B^{ij} h_i h_j a_i^* a_j^*) \right] \right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row. The IUCr Commission on Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: atom\_site.aniso\_B[1][2]\_esd (associated esd),

atom\_site.aniso\_U[1][2] (conversion constant),  
atom\_site\_anisotrop.U[1][2] (conversion constant),  
atom\_site.aniso\_U[1][2] (alternate exclusive),  
atom\_site\_anisotrop.B[1][2] (alternate exclusive),  
atom\_site\_anisotrop.U[1][2] (alternate exclusive).

[atom\_site]

atom\_site.aniso\_B[1][2]\_esd (float)

The standard uncertainty (estimated standard deviation) of atom\_site.aniso\_B[1][2].

Related items: atom\_site.aniso\_B[1][2] (associated value),  
atom\_site.aniso\_U[1][2]\_esd (conversion constant),  
atom\_site\_anisotrop.U[1][2]\_esd (conversion constant),  
atom\_site.aniso\_U[1][2]\_esd (alternate exclusive),  
atom\_site\_anisotrop.B[1][2]\_esd (alternate exclusive),  
atom\_site\_anisotrop.U[1][2]\_esd (alternate exclusive).

[atom\_site]

atom\_site.aniso\_B[1][3] (float, su)

The [1][3] element of the anisotropic atomic displacement matrix **B**, which appears in the structure-factor term as

$$T = \exp\left\{-\frac{1}{4} \sum_i \left[ \sum_j (B^{ij} h_i h_j a_i^* a_j^*) \right] \right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row. The IUCr Commission on

Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: `_atom_site.aniso_B[1][3]_esd` (associated esd),  
`_atom_site.aniso_U[1][3]` (conversion constant),  
`_atom_site_anisotrop.U[1][3]` (conversion constant),  
`_atom_site.aniso_U[1][3]` (alternate exclusive),  
`_atom_site_anisotrop.B[1][3]` (alternate exclusive),  
`_atom_site_anisotrop.U[1][3]` (alternate exclusive). [atom\_site]

`_atom_site.aniso_B[1][3]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site.aniso_B[1][3]`.

Related items: `_atom_site.aniso_B[1][3]` (associated value),  
`_atom_site.aniso_U[1][3]_esd` (conversion constant),  
`_atom_site_anisotrop.U[1][3]_esd` (conversion constant),  
`_atom_site.aniso_U[1][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[1][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[1][3]_esd` (alternate exclusive). [atom\_site]

`_atom_site.aniso_B[2][2]` (float, su)  
 The [2][2] element of the anisotropic atomic displacement matrix **B**, which appears in the structure-factor term as

$$T = \exp\left\{-\frac{1}{4} \sum_i \left[ \sum_j (B^{ij} h_i h_j a_i^* a_j^*) \right] \right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The IUCr Commission on Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: `_atom_site.aniso_B[2][2]_esd` (associated esd),  
`_atom_site.aniso_U[2][2]` (conversion constant),  
`_atom_site_anisotrop.U[2][2]` (conversion constant),  
`_atom_site.aniso_U[2][2]` (alternate exclusive),  
`_atom_site_anisotrop.B[2][2]` (alternate exclusive),  
`_atom_site_anisotrop.U[2][2]` (alternate exclusive). [atom\_site]

`_atom_site.aniso_B[2][2]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site.aniso_B[2][2]`.

Related items: `_atom_site.aniso_B[2][2]` (associated value),  
`_atom_site.aniso_U[2][2]_esd` (conversion constant),  
`_atom_site_anisotrop.U[2][2]_esd` (conversion constant),  
`_atom_site.aniso_U[2][2]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[2][2]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[2][2]_esd` (alternate exclusive). [atom\_site]

`_atom_site.aniso_B[2][3]` (float, su)  
 The [2][3] element of the anisotropic atomic displacement matrix **B**, which appears in the structure-factor term as

$$T = \exp\left\{-\frac{1}{4} \sum_i \left[ \sum_j (B^{ij} h_i h_j a_i^* a_j^*) \right] \right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row. The IUCr Commission on Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: `_atom_site.aniso_B[2][3]_esd` (associated esd),  
`_atom_site.aniso_U[2][3]` (conversion constant),  
`_atom_site_anisotrop.U[2][3]` (conversion constant),  
`_atom_site.aniso_U[2][3]` (alternate exclusive),  
`_atom_site_anisotrop.B[2][3]` (alternate exclusive),  
`_atom_site_anisotrop.U[2][3]` (alternate exclusive). [atom\_site]

`_atom_site.aniso_B[2][3]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site.aniso_B[2][3]`.

Related items: `_atom_site.aniso_B[2][3]` (associated value),  
`_atom_site.aniso_U[2][3]_esd` (conversion constant),  
`_atom_site_anisotrop.U[2][3]_esd` (conversion constant),  
`_atom_site.aniso_U[2][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[2][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[2][3]_esd` (alternate exclusive). [atom\_site]

`_atom_site.aniso_B[3][3]` (float, su)  
 The [3][3] element of the anisotropic atomic displacement matrix **B**, which appears in the structure-factor term as

$$T = \exp\left\{-\frac{1}{4} \sum_i \left[ \sum_j (B^{ij} h_i h_j a_i^* a_j^*) \right] \right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row. The IUCr Commission on Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: `_atom_site.aniso_B[3][3]_esd` (associated esd),  
`_atom_site.aniso_U[3][3]` (conversion constant),  
`_atom_site_anisotrop.U[3][3]` (conversion constant),  
`_atom_site.aniso_U[3][3]` (alternate exclusive),  
`_atom_site_anisotrop.B[3][3]` (alternate exclusive),  
`_atom_site_anisotrop.U[3][3]` (alternate exclusive). [atom\_site]

`_atom_site.aniso_B[3][3]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site.aniso_B[3][3]`.

Related items: `_atom_site.aniso_B[3][3]` (associated value),  
`_atom_site.aniso_U[3][3]_esd` (conversion constant),  
`_atom_site_anisotrop.U[3][3]_esd` (conversion constant),  
`_atom_site.aniso_U[3][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[3][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[3][3]_esd` (alternate exclusive). [atom\_site]

`_atom_site.aniso_ratio` (float)  
 Ratio of the maximum to minimum principal axes of displacement (thermal) ellipsoids.

The permitted range is [1.0, ∞).  
 Related item: `_atom_site_anisotrop.ratio` (alternate exclusive). [atom\_site]

`_atom_site.aniso_U[1][1]` (float, su)

The [1][1] element of the standard anisotropic atomic displacement matrix **U**, which appears in the structure-factor term as

$$T = \exp\left\{-2\pi^2 \sum_i \left[ \sum_j (U^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row.

Related items: `_atom_site.aniso_U[1][1]_esd` (associated esd),  
`_atom_site.aniso_B[1][1]` (conversion constant),  
`_atom_site_anisotrop.B[1][1]` (conversion constant),  
`_atom_site.aniso_B[1][1]` (alternate exclusive),  
`_atom_site_anisotrop.B[1][1]` (alternate exclusive),  
`_atom_site_anisotrop.U[1][1]` (alternate exclusive). [atom\_site]

`_atom_site.aniso_U[1][1]_esd` (float)

The standard uncertainty (estimated standard deviation) of `_atom_site.aniso_U[1][1]`.

Related items: `_atom_site.aniso_U[1][1]` (associated value),  
`_atom_site.aniso_B[1][1]_esd` (conversion constant),  
`_atom_site_anisotrop.B[1][1]_esd` (conversion constant),  
`_atom_site.aniso_B[1][1]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[1][1]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[1][1]_esd` (alternate exclusive). [atom\_site]

`_atom_site.aniso_U[1][2]` (float, su)

The [1][2] element of the standard anisotropic atomic displacement matrix **U**, which appears in the structure-factor term as

$$T = \exp\left\{-2\pi^2 \sum_i \left[ \sum_j (U^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row.

Related items: `_atom_site.aniso_U[1][2]_esd` (associated esd),  
`_atom_site.aniso_B[1][2]` (conversion constant),  
`_atom_site_anisotrop.B[1][2]` (conversion constant),  
`_atom_site.aniso_B[1][2]` (alternate exclusive),  
`_atom_site_anisotrop.B[1][2]` (alternate exclusive),  
`_atom_site_anisotrop.U[1][2]` (alternate exclusive). [atom\_site]

`_atom_site.aniso_U[1][2]_esd` (float)

The standard uncertainty (estimated standard deviation) of `_atom_site.aniso_U[1][2]`.

Related items: `_atom_site.aniso_U[1][2]` (associated value),  
`_atom_site.aniso_B[1][2]_esd` (conversion constant),  
`_atom_site_anisotrop.B[1][2]_esd` (conversion constant),  
`_atom_site.aniso_B[1][2]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[1][2]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[1][2]_esd` (alternate exclusive). [atom\_site]

`_atom_site.aniso_U[1][3]` (float, su)

The [1][3] element of the standard anisotropic atomic displacement matrix **U**, which appears in the structure-factor term as

$$T = \exp\left\{-2\pi^2 \sum_i \left[ \sum_j (U^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row.

Related items: `_atom_site.aniso_U[1][3]_esd` (associated esd),  
`_atom_site.aniso_B[1][3]` (conversion constant),  
`_atom_site_anisotrop.B[1][3]` (conversion constant),  
`_atom_site.aniso_B[1][3]` (alternate exclusive),  
`_atom_site_anisotrop.B[1][3]` (alternate exclusive),  
`_atom_site_anisotrop.U[1][3]` (alternate exclusive). [atom\_site]

`_atom_site.aniso_U[1][3]_esd` (float)

The standard uncertainty (estimated standard deviation) of `_atom_site.aniso_U[1][3]`.

Related items: `_atom_site.aniso_U[1][3]` (associated value),  
`_atom_site.aniso_B[1][3]_esd` (conversion constant),  
`_atom_site_anisotrop.B[1][3]_esd` (conversion constant),  
`_atom_site.aniso_B[1][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[1][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[1][3]_esd` (alternate exclusive). [atom\_site]

`_atom_site.aniso_U[2][2]` (float, su)

The [2][2] element of the standard anisotropic atomic displacement matrix **U**, which appears in the structure-factor term as

$$T = \exp\left\{-2\pi^2 \sum_i \left[ \sum_j (U^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row.

Related items: `_atom_site.aniso_U[2][2]_esd` (associated esd),  
`_atom_site.aniso_B[2][2]` (conversion constant),  
`_atom_site_anisotrop.B[2][2]` (conversion constant),  
`_atom_site.aniso_B[2][2]` (alternate exclusive),  
`_atom_site_anisotrop.B[2][2]` (alternate exclusive),  
`_atom_site_anisotrop.U[2][2]` (alternate exclusive). [atom\_site]

`_atom_site.aniso_U[2][2]_esd` (float)

The standard uncertainty (estimated standard deviation) of `_atom_site.aniso_U[2][2]`.

Related items: `_atom_site.aniso_U[2][2]` (associated value),  
`_atom_site.aniso_B[2][2]_esd` (conversion constant),  
`_atom_site_anisotrop.B[2][2]_esd` (conversion constant),  
`_atom_site.aniso_B[2][2]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[2][2]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[2][2]_esd` (alternate exclusive). [atom\_site]

**`_atom_site.aniso_U[2][3]`** *(float, su)*  
 The [2][3] element of the standard anisotropic atomic displacement matrix **U**, which appears in the structure-factor term as

$$T = \exp\left\{-2\pi^2 \sum_i \left[ \sum_j (U^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where *h* = the Miller indices and *a*\* = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either *B*'s or *U*'s, but not as both. The unique elements of the real symmetric matrix are entered by row.

Related items: `_atom_site.aniso_U[2][3]_esd` (associated esd),  
`_atom_site.aniso_B[2][3]` (conversion constant),  
`_atom_site_anisotrop.B[2][3]` (conversion constant),  
`_atom_site.aniso_B[2][3]` (alternate exclusive),  
`_atom_site_anisotrop.B[2][3]` (alternate exclusive),  
`_atom_site_anisotrop.U[2][3]` (alternate exclusive). [atom\_site]

**`_atom_site.aniso_U[2][3]_esd`** *(float)*  
 The standard uncertainty (estimated standard deviation) of `_atom_site.aniso_U[2][3]`.

Related items: `_atom_site.aniso_U[2][3]` (associated value),  
`_atom_site.aniso_B[2][3]_esd` (conversion constant),  
`_atom_site_anisotrop.B[2][3]_esd` (conversion constant),  
`_atom_site.aniso_B[2][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[2][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[2][3]_esd` (alternate exclusive). [atom\_site]

**`_atom_site.aniso_U[3][3]`** *(float, su)*  
 The [3][3] element of the standard anisotropic atomic displacement matrix **U**, which appears in the structure-factor term as

$$T = \exp\left\{-2\pi^2 \sum_i \left[ \sum_j (U^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where *h* = the Miller indices and *a*\* = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either *B*'s or *U*'s, but not as both. The unique elements of the real symmetric matrix are entered by row.

Related items: `_atom_site.aniso_U[3][3]_esd` (associated esd),  
`_atom_site.aniso_B[3][3]` (conversion constant),  
`_atom_site_anisotrop.B[3][3]` (conversion constant),  
`_atom_site.aniso_B[3][3]` (alternate exclusive),  
`_atom_site_anisotrop.B[3][3]` (alternate exclusive),  
`_atom_site_anisotrop.U[3][3]` (alternate exclusive). [atom\_site]

**`_atom_site.aniso_U[3][3]_esd`** *(float)*  
 The standard uncertainty (estimated standard deviation) of `_atom_site.aniso_U[3][3]`.

Related items: `_atom_site.aniso_U[3][3]` (associated value),  
`_atom_site.aniso_B[3][3]_esd` (conversion constant),  
`_atom_site_anisotrop.B[3][3]_esd` (conversion constant),  
`_atom_site.aniso_B[3][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[3][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[3][3]_esd` (alternate exclusive). [atom\_site]

**`_atom_site.attached_hydrogens`** *(int)*  
`_atom_site_attached_hydrogens` (*cif\_core.dic 2.0.1*)

The number of hydrogen atoms attached to the atom at this site excluding any hydrogen atoms for which coordinates (measured or calculated) are given.

The permitted range is [0, 8].

Examples: '2' (water oxygen), '1' (hydroxyl oxygen), '4' (ammonium nitrogen).

[atom\_site]

**`*_atom_site.auth_asym_id`** *(code)*

An alternative identifier for `_atom_site.label_asym_id` that may be provided by an author in order to match the identification used in the publication that describes the structure.

The following item(s) have an equivalent role in their respective categories:

`_geom_angle.atom_site_auth_asym_id_1`,  
`_geom_angle.atom_site_auth_asym_id_2`,  
`_geom_angle.atom_site_auth_asym_id_3`,  
`_geom_bond.atom_site_auth_asym_id_1`,  
`_geom_bond.atom_site_auth_asym_id_2`,  
`_geom_contact.atom_site_auth_asym_id_1`,  
`_geom_contact.atom_site_auth_asym_id_2`,  
`_geom_hbond.atom_site_auth_asym_id_A`,  
`_geom_hbond.atom_site_auth_asym_id_D`,  
`_geom_hbond.atom_site_auth_asym_id_H`,  
`_geom_torsion.atom_site_auth_asym_id_1`,  
`_geom_torsion.atom_site_auth_asym_id_2`,  
`_geom_torsion.atom_site_auth_asym_id_3`,  
`_geom_torsion.atom_site_auth_asym_id_4`,  
`_struct_conf.beg_auth_asym_id`,  
`_struct_conf.end_auth_asym_id`,  
`_struct_conn.ptnr1_auth_asym_id`,  
`_struct_conn.ptnr2_auth_asym_id`,  
`_struct_mon_nucl.auth_asym_id`,  
`_struct_mon_prot.auth_asym_id`,  
`_struct_mon_prot_cis.auth_asym_id`,  
`_struct_ncs_dom_lim.beg_auth_asym_id`,  
`_struct_ncs_dom_lim.end_auth_asym_id`,  
`_struct_sheet_range.beg_auth_asym_id`,  
`_struct_sheet_range.end_auth_asym_id`,  
`_struct_site_gen.auth_asym_id`. [atom\_site]

**`_atom_site.auth_atom_id`** *(atcode)*

An alternative identifier for `_atom_site.label_atom_id` that may be provided by an author in order to match the identification used in the publication that describes the structure.

The following item(s) have an equivalent role in their respective categories:

`_geom_angle.atom_site_auth_atom_id_1`,  
`_geom_angle.atom_site_auth_atom_id_2`,  
`_geom_angle.atom_site_auth_atom_id_3`,  
`_geom_bond.atom_site_auth_atom_id_1`,  
`_geom_bond.atom_site_auth_atom_id_2`,  
`_geom_contact.atom_site_auth_atom_id_1`,  
`_geom_contact.atom_site_auth_atom_id_2`,  
`_geom_hbond.atom_site_auth_atom_id_A`,  
`_geom_hbond.atom_site_auth_atom_id_D`,  
`_geom_hbond.atom_site_auth_atom_id_H`,  
`_geom_torsion.atom_site_auth_atom_id_1`,  
`_geom_torsion.atom_site_auth_atom_id_2`,  
`_geom_torsion.atom_site_auth_atom_id_3`,  
`_geom_torsion.atom_site_auth_atom_id_4`,  
`_struct_conn.ptnr1_auth_atom_id`,  
`_struct_conn.ptnr2_auth_atom_id`,  
`_struct_sheet_hbond.range_1_beg_auth_atom_id`,  
`_struct_sheet_hbond.range_1_end_auth_atom_id`,  
`_struct_sheet_hbond.range_2_beg_auth_atom_id`,  
`_struct_sheet_hbond.range_2_end_auth_atom_id`,  
`_struct_site_gen.auth_atom_id`. [atom\_site]

`_atom_site.auth_comp_id` (code)

An alternative identifier for `_atom_site.label_comp_id` that may be provided by an author in order to match the identification used in the publication that describes the structure.

The following item(s) have an equivalent role in their respective categories:

`_geom_angle.atom_site_auth_comp_id_1`,  
`_geom_angle.atom_site_auth_comp_id_2`,  
`_geom_angle.atom_site_auth_comp_id_3`,  
`_geom_bond.atom_site_auth_comp_id_1`,  
`_geom_bond.atom_site_auth_comp_id_2`,  
`_geom_contact.atom_site_auth_comp_id_1`,  
`_geom_contact.atom_site_auth_comp_id_2`,  
`_geom_hbond.atom_site_auth_comp_id_A`,  
`_geom_hbond.atom_site_auth_comp_id_D`,  
`_geom_hbond.atom_site_auth_comp_id_H`,  
`_geom_torsion.atom_site_auth_comp_id_1`,  
`_geom_torsion.atom_site_auth_comp_id_2`,  
`_geom_torsion.atom_site_auth_comp_id_3`,  
`_geom_torsion.atom_site_auth_comp_id_4`,  
`_struct_conf.beg_auth_comp_id`,  
`_struct_conf.end_auth_comp_id`,  
`_struct_conn.ptnr1_auth_comp_id`,  
`_struct_conn.ptnr2_auth_comp_id`,  
`_struct_mon_nucl.auth_comp_id`,  
`_struct_mon_prot.auth_comp_id`,  
`_struct_mon_prot_cis.auth_comp_id`,  
`_struct_ncs_dom_lim.beg_auth_comp_id`,  
`_struct_ncs_dom_lim.end_auth_comp_id`,  
`_struct_sheet_range.beg_auth_comp_id`,  
`_struct_sheet_range.end_auth_comp_id`,  
`_struct_site_gen.auth_comp_id`. [atom\_site]

`_atom_site.auth_seq_id` (code)

An alternative identifier for `_atom_site.label_seq_id` that may be provided by an author in order to match the identification used in the publication that describes the structure. Note that this is not necessarily a number, that the values do not have to be positive, and that the value does not have to correspond to the value of `_atom_site.label_seq_id`. The value of `_atom_site.label_seq_id` is required to be a sequential list of positive integers. The author may assign values to `_atom_site.auth_seq_id` in any desired way. For instance, the values may be used to relate this structure to a numbering scheme in a homologous structure, including sequence gaps or insertion codes. Alternatively, a scheme may be used for a truncated polymer that maintains the numbering scheme of the full length polymer. In all cases, the scheme used here must match the scheme used in the publication that describes the structure.

The following item(s) have an equivalent role in their respective categories:

`_geom_angle.atom_site_auth_seq_id_1`,  
`_geom_angle.atom_site_auth_seq_id_2`,  
`_geom_angle.atom_site_auth_seq_id_3`,  
`_geom_bond.atom_site_auth_seq_id_1`,  
`_geom_bond.atom_site_auth_seq_id_2`,  
`_geom_contact.atom_site_auth_seq_id_1`,  
`_geom_contact.atom_site_auth_seq_id_2`,  
`_geom_hbond.atom_site_auth_seq_id_A`,  
`_geom_hbond.atom_site_auth_seq_id_D`,  
`_geom_hbond.atom_site_auth_seq_id_H`,  
`_geom_torsion.atom_site_auth_seq_id_1`,  
`_geom_torsion.atom_site_auth_seq_id_2`,  
`_geom_torsion.atom_site_auth_seq_id_3`,  
`_geom_torsion.atom_site_auth_seq_id_4`,  
`_struct_conf.beg_auth_seq_id`,  
`_struct_conf.end_auth_seq_id`,  
`_struct_conn.ptnr1_auth_seq_id`,  
`_struct_conn.ptnr2_auth_seq_id`,  
`_struct_mon_nucl.auth_seq_id`.

`_struct_mon_prot.auth_seq_id`,  
`_struct_mon_prot_cis.auth_seq_id`,  
`_struct_ncs_dom_lim.beg_auth_seq_id`,  
`_struct_ncs_dom_lim.end_auth_seq_id`,  
`_struct_sheet_hbond.range_1_beg_auth_seq_id`,  
`_struct_sheet_hbond.range_1_end_auth_seq_id`,  
`_struct_sheet_hbond.range_2_beg_auth_seq_id`,  
`_struct_sheet_hbond.range_2_end_auth_seq_id`,  
`_struct_sheet_range.beg_auth_seq_id`,  
`_struct_sheet_range.end_auth_seq_id`,  
`_struct_site_gen.auth_seq_id`. [atom\_site]

`_atom_site.B_equiv_geom_mean` (float, su)

`_atom_site_B_equiv_geom_mean` (cif\_core.dic 2.0.1)

Equivalent isotropic atomic displacement parameter,  $B_{\text{eq}}$ , in ångströms squared, calculated as the geometric mean of the anisotropic atomic displacement parameters.

$$B_{\text{eq}} = (B_i B_j B_k)^{1/3},$$

where  $B_n$  = the principal components of the orthogonalized  $B^{ij}$ .

The IUCr Commission on Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

The permitted range is [0.0, ∞).

Related items: `_atom_site.B_equiv_geom_mean_esd` (associated esd),

`_atom_site.U_equiv_geom_mean` (conversion constant). [atom\_site]

`_atom_site.B_equiv_geom_mean_esd` (float)

The standard uncertainty (estimated standard deviation) of `_atom_site.B_equiv_geom_mean`.

Related items: `_atom_site.B_equiv_geom_mean` (associated value),

`_atom_site.U_equiv_geom_mean` (conversion constant). [atom\_site]

`_atom_site.B_iso_or_equiv` (float, su)

`_atom_site_B_iso_or_equiv` (cif\_core.dic 2.0.1)

Isotropic atomic displacement parameter, or equivalent isotropic atomic displacement parameter,  $B_{\text{eq}}$ , calculated from the anisotropic displacement parameters.

$$B_{\text{eq}} = (1/3) \sum_i \left[ \sum_j (B^{ij} A_i A_j a_i^* a_j^*) \right],$$

where  $A$  = the real-space cell lengths and  $a^*$  = the reciprocal-space cell lengths;  $B^{ij} = 8\pi^2 U^{ij}$ .

Reference: Fischer, R. X. & Tillmanns, E. (1988). *Acta Cryst.* **C44**, 775–776.

The IUCr Commission on Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: `_atom_site.B_iso_or_equiv_esd` (associated esd),

`_atom_site.U_iso_or_equiv` (conversion constant). [atom\_site]

`_atom_site.B_iso_or_equiv_esd` (float)

The standard uncertainty (estimated standard deviation) of `_atom_site.B_iso_or_equiv`.

Related items: `_atom_site.B_iso_or_equiv` (associated value),

`_atom_site.U_iso_or_equiv_esd` (conversion constant). [atom\_site]

`_atom_site.calc_attached_atom` (code)

`_atom_site_calc_attached_atom` (cif\_core.dic 2.0.1)

The `_atom_site.id` of the atom site to which the ‘geometry-calculated’ atom site is attached.

[atom\_site]

**\_atom\_site.calc\_flag** (ucode)  
**\_atom\_site\_calc\_flag** (cif\_core.dic 2.0.1)

A standard code to signal whether the site coordinates have been determined from the intensities or calculated from the geometry of surrounding sites, or have been assigned dummy values. The abbreviation 'c' may be used in place of 'calc'.

The data value must be one of the following:

d determined from experimental measurements  
 calc calculated from molecular geometry  
 c abbreviation for 'calc'  
 dum dummy site with meaningless coordinates

[atom\_site]

**\_atom\_site.Cartn\_x** (float, su)  
**\_atom\_site\_Cartn\_x** (cif\_core.dic 2.0.1)

The *x* atom-site coordinate in ångströms specified according to a set of orthogonal Cartesian axes related to the cell axes as specified by the description given in **\_atom\_sites.Cartn\_transform\_axes**.

Related item: **\_atom\_site.Cartn\_x\_esd** (associated esd).

[atom\_site]

**\_atom\_site.Cartn\_x\_esd** (float)

The standard uncertainty (estimated standard deviation) of **\_atom\_site.Cartn\_x**.

Related item: **\_atom\_site.Cartn\_x** (associated value).

[atom\_site]

**\_atom\_site.Cartn\_y** (float, su)  
**\_atom\_site\_Cartn\_y** (cif\_core.dic 2.0.1)

The *y* atom-site coordinate in ångströms specified according to a set of orthogonal Cartesian axes related to the cell axes as specified by the description given in **\_atom\_sites.Cartn\_transform\_axes**.

Related item: **\_atom\_site.Cartn\_y\_esd** (associated esd).

[atom\_site]

**\_atom\_site.Cartn\_y\_esd** (float)

The standard uncertainty (estimated standard deviation) of **\_atom\_site.Cartn\_y**.

Related item: **\_atom\_site.Cartn\_y** (associated value).

[atom\_site]

**\_atom\_site.Cartn\_z** (float, su)  
**\_atom\_site\_Cartn\_z** (cif\_core.dic 2.0.1)

The *z* atom-site coordinate in ångströms specified according to a set of orthogonal Cartesian axes related to the cell axes as specified by the description given in **\_atom\_sites.Cartn\_transform\_axes**.

Related item: **\_atom\_site.Cartn\_z\_esd** (associated esd).

[atom\_site]

**\_atom\_site.Cartn\_z\_esd** (float)

The standard uncertainty (estimated standard deviation) of **\_atom\_site.Cartn\_z**.

Related item: **\_atom\_site.Cartn\_z** (associated value).

[atom\_site]

**\_atom\_site.chemical\_conn\_number**  
**\_atom\_site\_chemical\_conn\_number** (cif\_core.dic 2.0.1)

This data item is a pointer to **\_chemical\_conn\_atom.number** in the CHEMICAL\_CONN\_ATOM category.

**\_atom\_site.constraints** (line)  
**\_atom\_site\_constraints** (cif\_core.dic 2.0.1)

A description of the constraints applied to parameters at this site during refinement. See also **\_atom\_site.refinement\_flags** and **\_refine.ls\_number\_constraints**.

Example: 'pop=1.0-pop (Zn3)'.

[atom\_site]

**\_atom\_site.details** (text)  
**\_atom\_site\_description** (cif\_core.dic 2.0.1)

A description of special aspects of this site. See also **\_atom\_site.refinement\_flags**.

Example: 'Ag/Si disordered'.

[atom\_site]

**\_atom\_site.disorder\_assembly** (code)  
**\_atom\_site\_disorder\_assembly** (cif\_core.dic 2.0.1)

A code which identifies a cluster of atoms that show long-range positional disorder but are locally ordered. Within each such cluster of atoms, **\_atom\_site.disorder\_group** is used to identify the sites that are simultaneously occupied. This field is only needed if there is more than one cluster of disordered atoms showing independent local order.

*Note:* This data item would not in general be used in a macromolecular data block.

[atom\_site]

**\_atom\_site.disorder\_group** (code)  
**\_atom\_site\_disorder\_group** (cif\_core.dic 2.0.1)

A code which identifies a group of positionally disordered atom sites that are locally simultaneously occupied. Atoms that are positionally disordered over two or more sites (e.g. the hydrogen atoms of a methyl group that exists in two orientations) can be assigned to two or more groups. Sites belonging to the same group are simultaneously occupied, but those belonging to different groups are not. A minus prefix (e.g. '-1') is used to indicate sites disordered about a special position.

*Note:* This data item would not in general be used in a macromolecular data block.

Where no value is given, the assumed value is '.'.

[atom\_site]

**\_atom\_site.footnote\_id**

The value of **\_atom\_site.footnote\_id** must match an ID specified by **\_atom\_sites\_footnote.id** in the ATOM\_SITES\_FOOTNOTE list.

**\_atom\_site.fract\_x** (float, su)  
**\_atom\_site\_fract\_x** (cif\_core.dic 2.0.1)

The *x* coordinate of the atom-site position specified as a fraction of **\_cell.length\_a**.

Related item: **\_atom\_site.fract\_x\_esd** (associated esd).

[atom\_site]

**\_atom\_site.fract\_x\_esd** (float)

The standard uncertainty (estimated standard deviation) of **\_atom\_site.fract\_x**.

Related item: **\_atom\_site.fract\_x** (associated value).

[atom\_site]

**\_atom\_site.fract\_y** (float, su)  
**\_atom\_site\_fract\_y** (cif\_core.dic 2.0.1)

The *y* coordinate of the atom-site position specified as a fraction of **\_cell.length\_b**.

Related item: **\_atom\_site.fract\_y\_esd** (associated esd).

[atom\_site]

**\_atom\_site.fract\_y\_esd** (float)

The standard uncertainty (estimated standard deviation) of **\_atom\_site.fract\_y**.

Related item: **\_atom\_site.fract\_y** (associated value).

[atom\_site]

**\_atom\_site.fract\_z** (float, su)  
**\_atom\_site\_fract\_z** (cif\_core.dic 2.0.1)

The *z* coordinate of the atom-site position specified as a fraction of **\_cell.length\_c**.

Related item: **\_atom\_site.fract\_z\_esd** (associated esd).

[atom\_site]

**atom\_site.fract\_z\_esd** (float)  
The standard uncertainty (estimated standard deviation) of atom\_site.fract\_z.  
Related item: atom\_site.fract\_z (associated value). [atom\_site]

**atom\_site.group\_PDB** (code)  
The group of atoms to which the atom site belongs. This data item is provided for compatibility with the original Protein Data Bank format, and only for that purpose.  
The data value must be one of the following:  
ATOM  
HETATM [atom\_site]

\* **atom\_site.id** (code)  
atom\_site\_label (cif\_core.dic 2.0.1)  
The value of atom\_site.id must uniquely identify a record in the ATOM\_SITE list. Note that this item need not be a number; it can be any unique identifier. This data item was introduced to provide compatibility between small-molecule and macromolecular CIFs. In a small-molecule CIF, atom\_site\_label is the identifier for the atom. In a macromolecular CIF, the atom identifier is the aggregate of atom\_site.label\_alt\_id, atom\_site.label\_asym\_id, atom\_site.label\_atom\_id, atom\_site.label\_comp\_id and atom\_site.label\_seq\_id. For the two types of files to be compatible, a formal identifier for the category had to be introduced that was independent of the different modes of identifying the atoms. For compatibility with older CIFs, atom\_site\_label is aliased to atom\_site.id.

The following item(s) have an equivalent role in their respective categories:

atom\_site\_anisotrop.id,  
geom\_angle.atom\_site\_id\_1,  
geom\_angle.atom\_site\_id\_2,  
geom\_angle.atom\_site\_id\_3,  
geom\_bond.atom\_site\_id\_1,  
geom\_bond.atom\_site\_id\_2,  
geom\_contact.atom\_site\_id\_1,  
geom\_contact.atom\_site\_id\_2,  
geom\_hbond.atom\_site\_id\_A,  
geom\_hbond.atom\_site\_id\_D,  
geom\_hbond.atom\_site\_id\_H,  
geom\_torsion.atom\_site\_id\_1,  
geom\_torsion.atom\_site\_id\_2,  
geom\_torsion.atom\_site\_id\_3,  
geom\_torsion.atom\_site\_id\_4.

Examples: 'S', 'C12', 'Ca3g28', 'Fe3+17', 'H\*251', 'boron2a', 'Ca.phe.83.a.0', 'Zn.Zn.301.A.0'. [atom\_site]

\* **atom\_site.label\_alt\_id**  
A component of the identifier for this atom site. For further details, see the definition of the ATOM\_SITE\_ALT category. This data item is a pointer to atom\_sites\_alt.id in the ATOM\_SITES\_ALT category.

\* **atom\_site.label\_asym\_id**  
A component of the identifier for this atom site. For further details, see the definition of the STRUCT\_ASYM category. This data item is a pointer to struct\_asym.id in the STRUCT\_ASYM category.

\* **atom\_site.label\_atom\_id**  
A component of the identifier for this atom site. This data item is a pointer to chem\_comp\_atom.atom\_id in the CHEM\_COMP\_ATOM category.

\* **atom\_site.label\_comp\_id**  
A component of the identifier for this atom site. This data item is a pointer to chem\_comp.id in the CHEM\_COMP category.

\* **atom\_site.label\_entity\_id**  
This data item is a pointer to entity.id in the ENTITY category.

\* **atom\_site.label\_seq\_id**  
This data item is a pointer to entity\_poly\_seq.num in the ENTITY\_POLY\_SEQ category.

**atom\_site.occupancy** (float, su)  
atom\_site\_occupancy (cif\_core.dic 2.0.1)  
The fraction of the atom type present at this site. The sum of the occupancies of all the atom types at this site may not significantly exceed 1.0 unless it is a dummy site.  
Related item: atom\_site.occupancy\_esd (associated esd). Where no value is given, the assumed value is '1.0'. [atom\_site]

**atom\_site.occupancy\_esd** (float)  
The standard uncertainty (estimated standard deviation) of atom\_site.occupancy.  
Related item: atom\_site.occupancy (associated value). [atom\_site]

**atom\_site.refinement\_flags** (code)  
atom\_site\_refinement\_flags (cif\_core.dic 2.3)  
A concatenated series of single-letter codes which indicate the refinement restraints or constraints applied to this site. This item should not be used. It has been replaced by atom\_site.refinement\_flags\_posn, \*\_adp and \*\_occupancy. It is retained in this dictionary only to provide compatibility with old CIFs.

Related items: atom\_site.refinement\_flags\_posn (replaces), atom\_site.refinement\_flags\_adp (replaces), atom\_site.refinement\_flags\_occupancy (replaces).

The data value must be one of the following:

. no refinement constraints  
S special-position constraint on site  
G rigid-group refinement of site  
R riding-atom site attached to non-riding atom  
D distance or angle restraint on site  
T thermal displacement constraints  
U  $U_{iso}$  or  $U^{ij}$  restraint (rigid bond)  
P partial occupancy constraint

[atom\_site]

**atom\_site.refinement\_flags\_adp** (code)  
atom\_site\_refinement\_flags\_adp (cif\_core.dic 2.3)  
A code which indicates the refinement restraints or constraints applied to the atomic displacement parameters of this site.

Related item: atom\_site.refinement\_flags\_posn (alternate).

The data value must be one of the following:

. no constraints on atomic displacement parameters  
T special-position constraints on atomic displacement parameters  
U  $U_{iso}$  or  $U^{ij}$  restraint (rigid bond)  
TU both constraints applied

[atom\_site]

**atom\_site.refinement\_flags\_occupancy** (code)  
atom\_site\_refinement\_flags\_occupancy (cif\_core.dic 2.3)  
A code which indicates that refinement restraints or constraints were applied to the occupancy of this site.

Related item: atom\_site.refinement\_flags\_posn (alternate).

The data value must be one of the following:

. no constraints on site-occupancy parameters  
P site-occupancy constraint

[atom\_site]

**\_atom\_site.refinement\_flags\_posn** (code)  
**\_atom\_site\_refinement\_flags\_posn** (cif\_core.dic 2.3)

A code which indicates the refinement restraints or constraints applied to the positional coordinates of this site.

Related item: **\_atom\_site.refinement\_flags\_posn** (alternate).

The data value must be one of the following:

- .
- D distance or angle restraint on positional coordinates
- G rigid-group refinement of positional coordinates
- R riding-atom site attached to non-riding atom
- S special-position constraint on positional coordinates
- DG combination of the above constraints
- DR combination of the above constraints
- DS combination of the above constraints
- GR combination of the above constraints
- GS combination of the above constraints
- RS combination of the above constraints
- DGR combination of the above constraints
- DGS combination of the above constraints
- DRS combination of the above constraints
- GRS combination of the above constraints
- DGRS combination of the above constraints

[atom\_site]

**\_atom\_site.restraints** (text)

**\_atom\_site\_restraints** (cif\_core.dic 2.0.1)

A description of restraints applied to specific parameters at this site during refinement. See also **\_atom\_site.refinement\_flags** and **\_refine.ls\_number\_restraints**.

Example: 'restrained to planar ring'.

[atom\_site]

**\_atom\_site.symmetry\_multiplicity** (int)

**\_atom\_site\_symmetry\_multiplicity** (cif\_core.dic 2.0.1)

The multiplicity of a site due to the space-group symmetry as is given in *International Tables for Crystallography* Vol. A (2002).

The permitted range is [1, 192].

[atom\_site]

**\_atom\_site.thermal\_displace\_type** (ucode)

**\_atom\_site\_thermal\_displace\_type** (cif\_core.dic 2.0.1)

A standard code used to describe the type of atomic displacement parameters used for the site.

The data value must be one of the following:

- Uani anisotropic  $U^{ij}$
- Uiso isotropic  $U$
- Uovl overall  $U$
- Umpe multipole expansion  $U$
- Bani anisotropic  $B^{ij}$
- Biso isotropic  $B$
- Bovl overall  $B$

[atom\_site]

\* **\_atom\_site.type\_symbol**

**\_atom\_site\_type\_symbol** (cif\_core.dic 2.0.1)

This data item is a pointer to **\_atom\_type.symbol** in the ATOM\_TYPE category.

**\_atom\_site.U\_equiv\_geom\_mean** (float, su)

**\_atom\_site\_U\_equiv\_geom\_mean** (cif\_core.dic 2.0.1)

Equivalent isotropic atomic displacement parameter,  $U_{eq}$ , in ångströms squared, calculated as the geometric mean of the anisotropic atomic displacement parameters.

$$U_{eq} = (U_i U_j U_k)^{1/3},$$

where  $U_n$  = the principal components of the orthogonalized  $U^{ij}$ .

The permitted range is [0.0, 10.0].

Related items: **\_atom\_site.U\_equiv\_geom\_mean\_esd** (associated esd),

**\_atom\_site.B\_equiv\_geom\_mean** (conversion constant). [atom\_site]

**\_atom\_site.U\_equiv\_geom\_mean\_esd** (float)

The standard uncertainty (estimated standard deviation) of **\_atom\_site.U\_equiv\_geom\_mean**.

Related items: **\_atom\_site.U\_equiv\_geom\_mean** (associated value),

**\_atom\_site.B\_equiv\_geom\_mean** (conversion constant). [atom\_site]

**\_atom\_site.U\_iso\_or\_equiv** (float, su)

**\_atom\_site\_U\_iso\_or\_equiv** (cif\_core.dic 2.0.1)

Isotropic atomic displacement parameter, or equivalent isotropic atomic displacement parameter,  $U_{eq}$ , calculated from anisotropic atomic displacement parameters.

$$U_{eq} = (1/3) \sum_i \left[ \sum_j (U^{ij} A_i A_j a_i^* a_j^*) \right],$$

where  $A$  = the real-space cell lengths and  $a^*$  = the reciprocal-space cell lengths.

Reference: Fischer, R. X. & Tillmanns, E. (1988). *Acta Cryst. C44*, 775–776.

The permitted range is [0.0, 10.0].

Related items: **\_atom\_site.U\_iso\_or\_equiv\_esd** (associated esd),

**\_atom\_site.B\_iso\_or\_equiv** (conversion constant). [atom\_site]

**\_atom\_site.U\_iso\_or\_equiv\_esd** (float)

The standard uncertainty (estimated standard deviation) of **\_atom\_site.U\_iso\_or\_equiv**.

Related items: **\_atom\_site.U\_iso\_or\_equiv** (associated value),

**\_atom\_site.B\_iso\_or\_equiv\_esd** (conversion constant). [atom\_site]

**\_atom\_site.Wyckoff\_symbol** (line)

**\_atom\_site\_Wyckoff\_symbol** (cif\_core.dic 2.0.1)

The Wyckoff symbol (letter) as listed in the space-group tables of *International Tables for Crystallography* Vol. A (2002).

[atom\_site]

### ATOM\_SITE\_ANISOTROP

Data items in the ATOM\_SITE\_ANISOTROP category record details about anisotropic displacement parameters. If the ATOM\_SITE\_ANISOTROP category is used for storing these data, the corresponding ATOM\_SITE data items are not used.

Category group(s): **inclusive\_group**

**atom\_group**

Category key(s): **\_atom\_site\_anisotrop.id**

*Example 1 – based on NDB structure BDL005 of Holbrook, Dickerson & Kim [Acta Cryst. (1985), B41, 255–262].*

loop\_

```

_atom_site_anisotrop.id
_atom_site_anisotrop.type_symbol
_atom_site_anisotrop.U[1][1]
_atom_site_anisotrop.U[1][2]
_atom_site_anisotrop.U[1][3]
_atom_site_anisotrop.U[2][2]
_atom_site_anisotrop.U[2][3]
_atom_site_anisotrop.U[3][3]

```

```

1 O  8642  4866  7299  -342  -258  -1427
2 C  5174  4871  6243  -1885 -2051  -1377
3 C  6202  5020  4395  -1130  -556  -632
4 O  4224  4700  5046  1105  -161  345
5 C  8684  4688  4171  -1850  -433  -292
6 O  11226  5255  3532  -341  2685  1328
7 C  10214  2428  5614  -2610 -1940  902
8 C  4590  3488  5827  751  -770  986
9 N  5014  4434  3447  -17  -1593  539
# ---- abbreviated ----

```

`_atom_site_anisotrop.B[1][1]` (float, su)  
`_atom_site_aniso_B_11` (cif\_core.dic 2.0.1)

The [1][1] element of the anisotropic atomic displacement matrix **B**, which appears in the structure-factor term as

$$T = \exp\left\{-\frac{1}{4} \sum_i \left[ \sum_j (B^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row. The IUCr Commission on Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: `_atom_site_anisotrop.B[1][1]_esd` (associated esd),  
`_atom_site.aniso_U[1][1]` (conversion constant),  
`_atom_site_anisotrop.U[1][1]` (conversion constant),  
`_atom_site.aniso_B[1][1]` (alternate exclusive),  
`_atom_site.aniso_U[1][1]` (alternate exclusive),  
`_atom_site_anisotrop.U[1][1]` (alternate exclusive).  
[atom\_site\_anisotrop]

`_atom_site_anisotrop.B[1][1]_esd` (float)  
The standard uncertainty (estimated standard deviation) of `_atom_site_anisotrop.B[1][1]`.

Related items: `_atom_site_anisotrop.B[1][1]` (associated value),  
`_atom_site.aniso_U[1][1]_esd` (conversion constant),  
`_atom_site_anisotrop.U[1][1]_esd` (conversion constant),  
`_atom_site.aniso_B[1][1]_esd` (alternate exclusive),  
`_atom_site.aniso_U[1][1]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[1][1]_esd` (alternate exclusive).  
[atom\_site\_anisotrop]

`_atom_site_anisotrop.B[1][2]` (float, su)  
`_atom_site_aniso_B_12` (cif\_core.dic 2.0.1)

The [1][2] element of the anisotropic atomic displacement matrix **B**, which appears in the structure-factor term as

$$T = \exp\left\{-\frac{1}{4} \sum_i \left[ \sum_j (B^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row. The IUCr Commission on Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: `_atom_site_anisotrop.B[1][2]_esd` (associated esd),  
`_atom_site.aniso_U[1][2]` (conversion constant),  
`_atom_site_anisotrop.U[1][2]` (conversion constant),  
`_atom_site.aniso_B[1][2]` (alternate exclusive),  
`_atom_site.aniso_U[1][2]` (alternate exclusive),  
`_atom_site_anisotrop.U[1][2]` (alternate exclusive).  
[atom\_site\_anisotrop]

`_atom_site_anisotrop.B[1][2]_esd` (float)  
The standard uncertainty (estimated standard deviation) of `_atom_site_anisotrop.B[1][2]`.

Related items: `_atom_site_anisotrop.B[1][2]` (associated value),  
`_atom_site.aniso_U[1][2]_esd` (conversion constant),  
`_atom_site_anisotrop.U[1][2]_esd` (conversion constant),  
`_atom_site.aniso_B[1][2]_esd` (alternate exclusive),  
`_atom_site.aniso_U[1][2]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[1][2]_esd` (alternate exclusive).  
[atom\_site\_anisotrop]

`_atom_site_anisotrop.B[1][3]` (float, su)  
`_atom_site_aniso_B_13` (cif\_core.dic 2.0.1)

The [1][3] element of the anisotropic atomic displacement matrix **B**, which appears in the structure-factor term as

$$T = \exp\left\{-\frac{1}{4} \sum_i \left[ \sum_j (B^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row. The IUCr Commission on Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: `_atom_site_anisotrop.B[1][3]_esd` (associated esd),  
`_atom_site.aniso_U[1][3]` (conversion constant),  
`_atom_site_anisotrop.U[1][3]` (conversion constant),  
`_atom_site.aniso_B[1][3]` (alternate exclusive),  
`_atom_site.aniso_U[1][3]` (alternate exclusive),  
`_atom_site_anisotrop.U[1][3]` (alternate exclusive).  
[atom\_site\_anisotrop]

`_atom_site_anisotrop.B[1][3]_esd` (float)  
The standard uncertainty (estimated standard deviation) of `_atom_site_anisotrop.B[1][3]`.

Related items: `_atom_site_anisotrop.B[1][3]` (associated value),  
`_atom_site.aniso_U[1][3]_esd` (conversion constant),  
`_atom_site_anisotrop.U[1][3]_esd` (conversion constant),  
`_atom_site.aniso_B[1][3]_esd` (alternate exclusive),  
`_atom_site.aniso_U[1][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[1][3]_esd` (alternate exclusive).  
[atom\_site\_anisotrop]

`_atom_site_anisotrop.B[2][2]` (float, su)  
`_atom_site_aniso_B_22` (cif\_core.dic 2.0.1)

The [2][2] element of the anisotropic atomic displacement matrix **B**, which appears in the structure-factor term as

$$T = \exp\left\{-\frac{1}{4} \sum_i \left[ \sum_j (B^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row. The IUCr Commission on Nomenclature recommends against the use of **B** for reporting

atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: `_atom_site_anisotrop.B[2][2]_esd` (associated esd),  
`_atom_site.aniso_U[2][2]` (conversion constant),  
`_atom_site_anisotrop.U[2][2]` (conversion constant),  
`_atom_site.aniso_B[2][2]` (alternate exclusive),  
`_atom_site.aniso_U[2][2]` (alternate exclusive),  
`_atom_site_anisotrop.U[2][2]` (alternate exclusive).  
[atom\_site\_anisotrop]

`_atom_site_anisotrop.B[2][2]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site_anisotrop.B[2][2]`.

Related items: `_atom_site_anisotrop.B[2][2]` (associated value),  
`_atom_site.aniso_U[2][2]_esd` (conversion constant),  
`_atom_site_anisotrop.U[2][2]_esd` (conversion constant),  
`_atom_site.aniso_B[2][2]_esd` (alternate exclusive),  
`_atom_site.aniso_U[2][2]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[2][2]_esd` (alternate exclusive).  
[atom\_site\_anisotrop]

`_atom_site_anisotrop.B[2][3]` (float, su)  
`_atom_site_aniso_B_23` (cif.core.dic 2.0.1)

The [2][3] element of the anisotropic atomic displacement matrix **B**, which appears in the structure-factor term as

$$T = \exp\left\{-\frac{1}{4} \sum_i \left[ \sum_j (B^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row. The IUCr Commission on Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: `_atom_site_anisotrop.B[2][3]_esd` (associated esd),  
`_atom_site.aniso_U[2][3]` (conversion constant),  
`_atom_site_anisotrop.U[2][3]` (conversion constant),  
`_atom_site.aniso_B[2][3]` (alternate exclusive),  
`_atom_site.aniso_U[2][3]` (alternate exclusive),  
`_atom_site_anisotrop.U[2][3]` (alternate exclusive).  
[atom\_site\_anisotrop]

`_atom_site_anisotrop.B[2][3]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site_anisotrop.B[2][3]`.

Related items: `_atom_site_anisotrop.B[2][3]` (associated value),  
`_atom_site.aniso_U[2][3]_esd` (conversion constant),  
`_atom_site_anisotrop.U[2][3]_esd` (conversion constant),  
`_atom_site.aniso_B[2][3]_esd` (alternate exclusive),  
`_atom_site.aniso_U[2][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[2][3]_esd` (alternate exclusive).  
[atom\_site\_anisotrop]

`_atom_site_anisotrop.B[3][3]` (float, su)  
`_atom_site_aniso_B_33` (cif.core.dic 2.0.1)

The [3][3] element of the anisotropic atomic displacement matrix **B**, which appears in the structure-factor term as

$$T = \exp\left\{-\frac{1}{4} \sum_i \left[ \sum_j (B^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row. The IUCr Commission on Nomenclature recommends against the use of **B** for reporting atomic displacement parameters. **U**, being directly proportional to **B**, is preferred.

Related items: `_atom_site_anisotrop.B[3][3]_esd` (associated esd),  
`_atom_site.aniso_U[3][3]` (conversion constant),  
`_atom_site_anisotrop.U[3][3]` (conversion constant),  
`_atom_site.aniso_B[3][3]` (alternate exclusive),  
`_atom_site.aniso_U[3][3]` (alternate exclusive),  
`_atom_site_anisotrop.U[3][3]` (alternate exclusive).  
[atom\_site\_anisotrop]

`_atom_site_anisotrop.B[3][3]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site_anisotrop.B[3][3]`.

Related items: `_atom_site_anisotrop.B[3][3]` (associated value),  
`_atom_site.aniso_U[3][3]_esd` (conversion constant),  
`_atom_site_anisotrop.U[3][3]_esd` (conversion constant),  
`_atom_site.aniso_B[3][3]_esd` (alternate exclusive),  
`_atom_site.aniso_U[3][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.U[3][3]_esd` (alternate exclusive).  
[atom\_site\_anisotrop]

\* `_atom_site_anisotrop.id`  
`_atom_site_aniso_label` (cif.core.dic 2.0.1)  
 This data item is a pointer to `_atom_site.id` in the ATOM\_SITE category.

`_atom_site_anisotrop.ratio` (float)  
`_atom_site_aniso_ratio` (cif.core.dic 2.0.1)  
 Ratio of the maximum to minimum principal axes of displacement (thermal) ellipsoids.  
 The permitted range is [1.0, ∞).  
 Related item: `_atom_site.aniso_ratio` (alternate exclusive).  
[atom\_site\_anisotrop]

\* `_atom_site_anisotrop.type_symbol`  
`_atom_site_aniso_type_symbol` (cif.core.dic 2.0.1)  
 This data item is a pointer to `_atom_type.symbol` in the ATOM\_TYPE category.

`_atom_site_anisotrop.U[1][1]` (float, su)  
`_atom_site_aniso_U_11` (cif.core.dic 2.0.1)  
 The [1][1] element of the standard anisotropic atomic displacement matrix **U**, which appears in the structure-factor term as

$$T = \exp\left\{-2\pi^2 \sum_i \left[ \sum_j (U^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where  $h$  = the Miller indices and  $a^*$  = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either  $B$ 's or  $U$ 's, but not as both. The unique elements of the real symmetric matrix are entered by row.

Related items: `_atom_site_anisotrop.U[1][1]_esd` (associated esd),  
`_atom_site.aniso_B[1][1]` (conversion constant),  
`_atom_site_anisotrop.B[1][1]` (conversion constant),  
`_atom_site.aniso_B[1][1]` (alternate exclusive),  
`_atom_site.aniso_U[1][1]` (alternate exclusive),  
`_atom_site_anisotrop.B[1][1]` (alternate exclusive).  
[atom\_site\_anisotrop]

`_atom_site_anisotrop.U[1][1]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site_anisotrop.U[1][1]`.

Related items: `_atom_site_anisotrop.U[1][1]` (associated value),  
`_atom_site.aniso_B[1][1]_esd` (conversion constant),  
`_atom_site_anisotrop.B[1][1]_esd` (conversion constant),  
`_atom_site.aniso_B[1][1]_esd` (alternate exclusive),  
`_atom_site.aniso_U[1][1]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[1][1]_esd` (alternate exclusive).

[atom\_site\_anisotrop]

`_atom_site_anisotrop.U[1][2]` (float, su)  
`_atom_site_aniso_U_12` (cif.core.dic 2.0.1)

The [1][2] element of the standard anisotropic atomic displacement matrix **U**, which appears in the structure-factor term as

$$T = \exp\left\{-2\pi^2 \sum_i \left[ \sum_j (U^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where *h* = the Miller indices and *a\** = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either *B*'s or *U*'s, but not as both. The unique elements of the real symmetric matrix are entered by row.

Related items: `_atom_site_anisotrop.U[1][2]_esd` (associated esd),  
`_atom_site.aniso_B[1][2]` (conversion constant),  
`_atom_site_anisotrop.B[1][2]` (conversion constant),  
`_atom_site.aniso_B[1][2]` (alternate exclusive),  
`_atom_site.aniso_U[1][2]` (alternate exclusive),  
`_atom_site_anisotrop.B[1][2]` (alternate exclusive).

[atom\_site\_anisotrop]

`_atom_site_anisotrop.U[1][2]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site_anisotrop.U[1][2]`.

Related items: `_atom_site_anisotrop.U[1][2]` (associated value),  
`_atom_site.aniso_B[1][2]_esd` (conversion constant),  
`_atom_site_anisotrop.B[1][2]_esd` (conversion constant),  
`_atom_site.aniso_B[1][2]_esd` (alternate exclusive),  
`_atom_site.aniso_U[1][2]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[1][2]_esd` (alternate exclusive).

[atom\_site\_anisotrop]

`_atom_site_anisotrop.U[1][3]` (float, su)  
`_atom_site_aniso_U_13` (cif.core.dic 2.0.1)

The [1][3] element of the standard anisotropic atomic displacement matrix **U**, which appears in the structure-factor term as

$$T = \exp\left\{-2\pi^2 \sum_i \left[ \sum_j (U^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where *h* = the Miller indices and *a\** = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either *B*'s or *U*'s, but not as both. The unique elements of the real symmetric matrix are entered by row.

Related items: `_atom_site_anisotrop.U[1][3]_esd` (associated esd),  
`_atom_site.aniso_B[1][3]` (conversion constant),  
`_atom_site_anisotrop.B[1][3]` (conversion constant),  
`_atom_site.aniso_B[1][3]` (alternate exclusive),  
`_atom_site.aniso_U[1][3]` (alternate exclusive),  
`_atom_site_anisotrop.B[1][3]` (alternate exclusive).

[atom\_site\_anisotrop]

`_atom_site_anisotrop.U[1][3]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site_anisotrop.U[1][3]`.

Related items: `_atom_site_anisotrop.U[1][3]` (associated value),  
`_atom_site.aniso_B[1][3]_esd` (conversion constant),  
`_atom_site_anisotrop.B[1][3]_esd` (conversion constant),  
`_atom_site.aniso_B[1][3]_esd` (alternate exclusive),  
`_atom_site.aniso_U[1][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[1][3]_esd` (alternate exclusive).

[atom\_site\_anisotrop]

`_atom_site_anisotrop.U[2][2]` (float, su)  
`_atom_site_aniso_U_22` (cif.core.dic 2.0.1)

The [2][2] element of the standard anisotropic atomic displacement matrix **U**, which appears in the structure-factor term as

$$T = \exp\left\{-2\pi^2 \sum_i \left[ \sum_j (U^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where *h* = the Miller indices and *a\** = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either *B*'s or *U*'s, but not as both. The unique elements of the real symmetric matrix are entered by row.

Related items: `_atom_site_anisotrop.U[2][2]_esd` (associated esd),  
`_atom_site.aniso_B[2][2]` (conversion constant),  
`_atom_site_anisotrop.B[2][2]` (conversion constant),  
`_atom_site.aniso_B[2][2]` (alternate exclusive),  
`_atom_site.aniso_U[2][2]` (alternate exclusive),  
`_atom_site_anisotrop.B[2][2]` (alternate exclusive).

[atom\_site\_anisotrop]

`_atom_site_anisotrop.U[2][2]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site_anisotrop.U[2][2]`.

Related items: `_atom_site_anisotrop.U[2][2]` (associated value),  
`_atom_site.aniso_B[2][2]_esd` (conversion constant),  
`_atom_site_anisotrop.B[2][2]_esd` (conversion constant),  
`_atom_site.aniso_B[2][2]_esd` (alternate exclusive),  
`_atom_site.aniso_U[2][2]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[2][2]_esd` (alternate exclusive).

[atom\_site\_anisotrop]

`_atom_site_anisotrop.U[2][3]` (float, su)  
`_atom_site_aniso_U_23` (cif.core.dic 2.0.1)

The [2][3] element of the standard anisotropic atomic displacement matrix **U**, which appears in the structure-factor term as

$$T = \exp\left\{-2\pi^2 \sum_i \left[ \sum_j (U^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where *h* = the Miller indices and *a\** = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either *B*'s or *U*'s, but not as both. The unique elements of the real symmetric matrix are entered by row.

Related items: `_atom_site_anisotrop.U[2][3]_esd` (associated esd),  
`_atom_site.aniso_B[2][3]` (conversion constant),  
`_atom_site_anisotrop.B[2][3]` (conversion constant),  
`_atom_site.aniso_B[2][3]` (alternate exclusive),  
`_atom_site.aniso_U[2][3]` (alternate exclusive),  
`_atom_site_anisotrop.B[2][3]` (alternate exclusive).

[atom\_site\_anisotrop]

`_atom_site_anisotrop.U[2][3]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site_anisotrop.U[2][3]`.

Related items: `_atom_site_anisotrop.U[2][3]` (associated value),  
`_atom_site.aniso_B[2][3]_esd` (conversion constant),  
`_atom_site_anisotrop.B[2][3]_esd` (conversion constant),  
`_atom_site.aniso_B[2][3]_esd` (alternate exclusive),  
`_atom_site.aniso_U[2][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[2][3]_esd` (alternate exclusive).

[atom\_site\_anisotrop]

`_atom_site_anisotrop.U[3][3]` (float, su)  
`_atom_site_aniso_U_33` (cif-core.dic 2.0.1)

The [3][3] element of the standard anisotropic atomic displacement matrix **U**, which appears in the structure-factor term as

$$T = \exp\left\{-2\pi^2 \sum_i \left[ \sum_j (U^{ij} h_i h_j a_i^* a_j^*) \right]\right\},$$

where *h* = the Miller indices and *a*<sup>\*</sup> = the reciprocal-space cell lengths.

These matrix elements may appear with atomic coordinates in the ATOM\_SITE category, or they may appear in the separate ATOM\_SITE\_ANISOTROP category, but they may not appear in both places. Similarly, anisotropic displacements may appear as either *B*'s or *U*'s, but not as both. The unique elements of the real symmetric matrix are entered by row.

Related items: `_atom_site_anisotrop.U[3][3]_esd` (associated esd),  
`_atom_site.aniso_B[3][3]` (conversion constant),  
`_atom_site_anisotrop.B[3][3]` (conversion constant),  
`_atom_site.aniso_B[3][3]` (alternate exclusive),  
`_atom_site.aniso_U[3][3]` (alternate exclusive),  
`_atom_site_anisotrop.B[3][3]` (alternate exclusive).

[atom\_site\_anisotrop]

`_atom_site_anisotrop.U[3][3]_esd` (float)  
 The standard uncertainty (estimated standard deviation) of `_atom_site_anisotrop.U[3][3]`.

Related items: `_atom_site_anisotrop.U[3][3]` (associated value),  
`_atom_site.aniso_B[3][3]_esd` (conversion constant),  
`_atom_site_anisotrop.B[3][3]_esd` (conversion constant),  
`_atom_site.aniso_B[3][3]_esd` (alternate exclusive),  
`_atom_site.aniso_U[3][3]_esd` (alternate exclusive),  
`_atom_site_anisotrop.B[3][3]_esd` (alternate exclusive).

[atom\_site\_anisotrop]

### ATOM\_SITES

Data items in the ATOM\_SITES category record details about the crystallographic cell and cell transformations, which are common to all atom sites.

Category group(s): `inclusive_group`  
`atom_group`

Category key(s): `_atom_sites.entry_id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```

_atom_sites.entry_id          '5HVP'
_atom_sites.Cartn_transform_axes
  'c along z, a* along x, b along y'
_atom_sites.Cartn_transf_matrix[1][1] 58.39
_atom_sites.Cartn_transf_matrix[1][2] 0.00
_atom_sites.Cartn_transf_matrix[1][3] 0.00
_atom_sites.Cartn_transf_matrix[2][1] 0.00
_atom_sites.Cartn_transf_matrix[2][2] 86.70
_atom_sites.Cartn_transf_matrix[2][3] 0.00
_atom_sites.Cartn_transf_matrix[3][1] 0.00
_atom_sites.Cartn_transf_matrix[3][2] 0.00
_atom_sites.Cartn_transf_matrix[3][3] 46.27
_atom_sites.Cartn_transf_vector[1] 0.00
_atom_sites.Cartn_transf_vector[2] 0.00
_atom_sites.Cartn_transf_vector[3] 0.00
    
```

`_atom_sites.Cartn_transf_matrix[1][1]` (float)  
`_atom_sites.Cartn_tran_matrix_11` (cif-core.dic 2.0.1)

The [1][1] element of the 3 × 3 matrix used to transform fractional coordinates in the ATOM\_SITE category to Cartesian coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The 3 × 1 translation is defined in `_atom_sites.Cartn_transf_vector[]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{fractional}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.Cartn_transf_matrix[1][2]` (float)  
`_atom_sites.Cartn_tran_matrix_12` (cif-core.dic 2.0.1)

The [1][2] element of the 3 × 3 matrix used to transform fractional coordinates in the ATOM\_SITE category to Cartesian coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The 3 × 1 translation is defined in `_atom_sites.Cartn_transf_vector[]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{fractional}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.Cartn_transf_matrix[1][3]` (float)  
`_atom_sites.Cartn_tran_matrix_13` (cif-core.dic 2.0.1)

The [1][3] element of the 3 × 3 matrix used to transform fractional coordinates in the ATOM\_SITE category to Cartesian coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The 3 × 1 translation is defined in `_atom_sites.Cartn_transf_vector[]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{fractional}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.Cartn_transf_matrix[2][1]` (float)  
`_atom_sites.Cartn_tran_matrix_21` (cif-core.dic 2.0.1)

The [2][1] element of the 3 × 3 matrix used to transform fractional coordinates in the ATOM\_SITE category to Cartesian coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The 3 × 1 translation is defined in `_atom_sites.Cartn_transf_vector[]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{fractional}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.Cartn_transf_matrix[2][2]` (float)  
`_atom_sites.Cartn_tran_matrix_22` (cif\_core.dic 2.0.1)

The [2][2] element of the  $3 \times 3$  matrix used to transform fractional coordinates in the ATOM\_SITE category to Cartesian coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The  $3 \times 1$  translation is defined in `_atom_sites.Cartn_transf_vector[]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{fractional}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.Cartn_transf_matrix[2][3]` (float)  
`_atom_sites.Cartn_tran_matrix_23` (cif\_core.dic 2.0.1)

The [2][3] element of the  $3 \times 3$  matrix used to transform fractional coordinates in the ATOM\_SITE category to Cartesian coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The  $3 \times 1$  translation is defined in `_atom_sites.Cartn_transf_vector[]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{fractional}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.Cartn_transf_matrix[3][1]` (float)  
`_atom_sites.Cartn_tran_matrix_31` (cif\_core.dic 2.0.1)

The [3][1] element of the  $3 \times 3$  matrix used to transform fractional coordinates in the ATOM\_SITE category to Cartesian coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The  $3 \times 1$  translation is defined in `_atom_sites.Cartn_transf_vector[]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{fractional}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.Cartn_transf_matrix[3][2]` (float)  
`_atom_sites.Cartn_tran_matrix_32` (cif\_core.dic 2.0.1)

The [3][2] element of the  $3 \times 3$  matrix used to transform fractional coordinates in the ATOM\_SITE category to Cartesian coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The  $3 \times 1$  translation is defined in `_atom_sites.Cartn_transf_vector[]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{fractional}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.Cartn_transf_matrix[3][3]` (float)  
`_atom_sites.Cartn_tran_matrix_33` (cif\_core.dic 2.0.1)

The [3][3] element of the  $3 \times 3$  matrix used to transform fractional coordinates in the ATOM\_SITE category to Cartesian coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The  $3 \times 1$  translation is defined in `_atom_sites.Cartn_transf_vector[]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{fractional}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.Cartn_transf_vector[1]` (float)  
`_atom_sites.Cartn_tran_vector_1` (cif\_core.dic 2.0.1)

The [1] element of the three-element vector used to transform fractional coordinates in the ATOM\_SITE category to Cartesian coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The rotation matrix is defined in `_atom_sites.Cartn_transf_matrix[][]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{fractional}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.Cartn_transf_vector[2]` (float)  
`_atom_sites.Cartn_tran_vector_2` (cif\_core.dic 2.0.1)

The [2] element of the three-element vector used to transform fractional coordinates in the ATOM\_SITE category to Cartesian coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The rotation matrix is defined in `_atom_sites.Cartn_transf_matrix[][]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{fractional}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.Cartn_transf_vector[3]` (float)  
`_atom_sites.Cartn_tran_vector_3` (cif\_core.dic 2.0.1)

The [3] element of the three-element vector used to transform fractional coordinates in the ATOM\_SITE category to Cartesian coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The rotation matrix is defined in `_atom_sites.Cartn_transf_matrix[][]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{fractional}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.Cartn_transform_axes` (text)  
`_atom_sites.Cartn_transform_axes` (cif\_core.dic 2.0.1)

A description of the relative alignment of the crystal cell axes to the Cartesian orthogonal axes as applied in the transformation matrix `_atom_sites.Cartn_transf_matrix[][]`.

Example: 'a parallel to x; b in the plane of y and z'.

[atom\_sites]

\* `_atom_sites.entry_id`

This data item is a pointer to `_entry.id` in the ENTRY category.



`_atom_sites.fract_transf_vector[1]` (float)  
`_atom_sites_fract_tran_vector_1` (cif\_core.dic 2.0.1)

The [1] element of the three-element vector used to transform Cartesian coordinates in the ATOM\_SITE category to fractional coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The  $3 \times 3$  rotation is defined in `_atom_sites.fract_transf_matrix[][]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{fractional}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.fract_transf_vector[2]` (float)  
`_atom_sites_fract_tran_vector_2` (cif\_core.dic 2.0.1)

The [2] element of the three-element vector used to transform Cartesian coordinates in the ATOM\_SITE category to fractional coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The  $3 \times 3$  rotation is defined in `_atom_sites.fract_transf_matrix[][]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{fractional}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.fract_transf_vector[3]` (float)  
`_atom_sites_fract_tran_vector_3` (cif\_core.dic 2.0.1)

The [3] element of the three-element vector used to transform Cartesian coordinates in the ATOM\_SITE category to fractional coordinates in the same category. The axial alignments of this transformation are described in `_atom_sites.Cartn_transform_axes`. The  $3 \times 3$  rotation is defined in `_atom_sites.fract_transf_matrix[][]`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{fractional}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} + \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

[atom\_sites]

`_atom_sites.solution_hydrogens` (ucode)  
`_atom_sites_solution_hydrogens` (cif\_core.dic 2.0.1)

This code identifies the method used to locate the hydrogen atoms.  
*Note:* This data item would not in general be used in a macro-molecular data block.

The data value must be one of the following:

difmap difference Fourier map  
 vecmap real-space vector search  
 heavy heavy-atom method  
 direct structure-invariant direct methods  
 geom inferred from neighbouring sites  
 disper anomalous-dispersion techniques  
 isomor isomorphous structure methods

[atom\_sites]

`_atom_sites.solution_primary` (ucode)  
`_atom_sites_solution_primary` (cif\_core.dic 2.0.1)

This code identifies the method used to locate the initial atom sites.

*Note:* This data item would not in general be used in a macro-molecular data block.

The data value must be one of the following:

difmap difference Fourier map  
 vecmap real-space vector search  
 heavy heavy-atom method  
 direct structure-invariant direct methods  
 geom inferred from neighbouring sites  
 disper anomalous-dispersion techniques  
 isomor isomorphous structure methods

[atom\_sites]

`_atom_sites.solution_secondary` (ucode)  
`_atom_sites_solution_secondary` (cif\_core.dic 2.0.1)

This code identifies the method used to locate the non-hydrogen-atom sites not found by `_atom_sites.solution_primary`.

*Note:* This data item would not in general be used in a macro-molecular data block.

The data value must be one of the following:

difmap difference Fourier map  
 vecmap real-space vector search  
 heavy heavy-atom method  
 direct structure-invariant direct methods  
 geom inferred from neighbouring sites  
 disper anomalous-dispersion techniques  
 isomor isomorphous structure methods

[atom\_sites]

`_atom_sites.special_details` (text)  
`_atom_sites_special_details` (cif\_core.dic 2.3)

Additional information about the atomic coordinates not coded elsewhere in the CIF.

[atom\_sites]

### ATOM\_SITES\_ALT

Data items in the ATOM\_SITES\_ALT category record details about the structural ensembles that should be generated from atom sites or groups of atom sites that are modelled in alternative conformations in this data block.

Category group(s): `inclusive_group`

`atom_group`

Category key(s): `_atom_sites_alt.id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _atom_sites_alt.id
  _atom_sites_alt.details
  .
  ; Atom sites with the alternative ID set to null are not
  modeled in alternative conformations
  ;
  1
  ; Atom sites with the alternative ID set to 1 have been
  modeled in alternative conformations with respect to atom
  sites marked with alternative ID 2. The conformations of
  amino-acid side chains and solvent atoms with alternative
  ID set to 1 correlate with the conformation of the
  inhibitor marked with alternative ID 1. They have been
  given an occupancy of 0.58 to match the occupancy assigned
  to the inhibitor.
  ;
```

```

2
; Atom sites with the alternative ID set to 2 have been
modeled in alternative conformations with respect to atom
sites marked with alternative ID 1. The conformations of
amino-acid side chains and solvent atoms with alternative
ID set to 2 correlate with the conformation of the
inhibitor marked with alternative ID 2. They have been
given an occupancy of 0.42 to match the occupancy assigned
to the inhibitor.
;
3
; Atom sites with the alternative ID set to 3 have been
modeled in alternative conformations with respect to
atoms marked with alternative ID 4. The conformations of
amino-acid side chains and solvent atoms with alternative
ID set to 3 do not correlate with the conformation of the
inhibitor. These atom sites have arbitrarily been given
an occupancy of 0.50.
;
4
; Atom sites with the alternative ID set to 4 have been
modeled in alternative conformations with respect to
atoms marked with alternative ID 3. The conformations of
amino-acid side chains and solvent atoms with alternative
ID set to 4 do not correlate with the conformation of the
inhibitor. These atom sites have arbitrarily been given
an occupancy of 0.50.
;

```

`_atom_sites_alt.details` (text)

A description of special aspects of the modelling of atoms in alternative conformations.

[atom\_sites\_alt]

\*`_atom_sites_alt.id` (code)

The value of `_atom_sites_alt.id` must uniquely identify a record in the ATOM\_SITES\_ALT list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

```

_atom_site.label_alt_id,
_atom_sites_alt_gen.alt_id,
_geom_angle.atom_site_label_alt_id_1,
_geom_angle.atom_site_label_alt_id_2,
_geom_angle.atom_site_label_alt_id_3,
_geom_bond.atom_site_label_alt_id_1,
_geom_bond.atom_site_label_alt_id_2,
_geom_contact.atom_site_label_alt_id_1,
_geom_contact.atom_site_label_alt_id_2,
_geom_hbond.atom_site_label_alt_id_A,
_geom_hbond.atom_site_label_alt_id_D,
_geom_hbond.atom_site_label_alt_id_H,
_geom_torsion.atom_site_label_alt_id_1,
_geom_torsion.atom_site_label_alt_id_2,
_geom_torsion.atom_site_label_alt_id_3,
_geom_torsion.atom_site_label_alt_id_4,
_struct_conn.ptnr1_label_alt_id,
_struct_conn.ptnr2_label_alt_id,
_struct_mon_nucl.label_alt_id,
_struct_mon_prot.label_alt_id,
_struct_mon_prot_cis.label_alt_id,
_struct_ncs_dom_lim.beg_label_alt_id,
_struct_ncs_dom_lim.end_label_alt_id,
_struct_site_gen.label_alt_id.

```

Examples: 'orientation 1', 'molecule abc'.

[atom\_sites\_alt]

## ATOM\_SITES\_ALT\_ENS

Data items in the ATOM\_SITES\_ALT\_ENS category record details about the ensemble structure generated from atoms with various alternative conformation IDs.

Category group(s): `inclusive_group`  
`atom_group`

Category key(s): `_atom_sites_alt_ens.id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```

loop_
_atom_sites_alt_ens.id
_atom_sites_alt_ens.details
'Ensemble 1-A'
; The inhibitor binds to the enzyme in two, roughly twofold
symmetric alternative conformations.

```

This conformational ensemble includes the more populated conformation of the inhibitor (ID=1) and the amino-acid side chains and solvent structure that correlate with this inhibitor conformation.

Also included are one set (ID=3) of side chains with alternative conformations when the conformations are not correlated with the inhibitor conformation.

```

;
'Ensemble 1-B'
; The inhibitor binds to the enzyme in two, roughly twofold
symmetric alternative conformations.

```

This conformational ensemble includes the more populated conformation of the inhibitor (ID=1) and the amino-acid side chains and solvent structure that correlate with this inhibitor conformation.

Also included are one set (ID=4) of side chains with alternative conformations when the conformations are not correlated with the inhibitor conformation.

```

;
'Ensemble 2-A'
; The inhibitor binds to the enzyme in two, roughly twofold
symmetric alternative conformations.

```

This conformational ensemble includes the less populated conformation of the inhibitor (ID=2) and the amino-acid side chains and solvent structure that correlate with this inhibitor conformation.

Also included are one set (ID=3) of side chains with alternative conformations when the conformations are not correlated with the inhibitor conformation.

```

;
'Ensemble 2-B'
; The inhibitor binds to the enzyme in two, roughly twofold
symmetric alternative conformations.

```

This conformational ensemble includes the less populated conformation of the inhibitor (ID=2) and the amino-acid side chains and solvent structure that correlate with this inhibitor conformation.

Also included are one set (ID=4) of side chains with alternative conformations when the conformations are not correlated with the inhibitor conformation.

```

;

```

`_atom_sites_alt_ens.details` (text)

A description of special aspects of the ensemble structure generated from atoms with various alternative IDs.

[atom\_sites\_alt\_ens]

\*`_atom_sites_alt_ens.id` (code)

The value of `_atom_sites_alt_ens.id` must uniquely identify a record in the ATOM\_SITES\_ALT\_ENS list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

```

_atom_sites_alt_gen.ens_id. [atom_sites_alt_ens]

```

**ATOM\_SITES\_ALT\_GEN**

Data items in the ATOM\_SITES\_ALT\_GEN category record details about the interpretation of multiple conformations in the structure.

Category group(s): **inclusive\_group**  
**atom\_group**

Category key(s): **\_atom\_sites\_alt\_gen.ens\_id**  
**\_atom\_sites\_alt\_gen.alt\_id**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _atom_sites_alt_gen.ens_id
  _atom_sites_alt_gen.alt_id
  'Ensemble 1-A' .
  'Ensemble 1-A' 1
  'Ensemble 1-A' 3
  'Ensemble 1-B' .
  'Ensemble 1-B' 1
  'Ensemble 1-B' 4
  'Ensemble 2-A' .
  'Ensemble 2-A' 2
  'Ensemble 2-A' 3
  'Ensemble 2-B' .
  'Ensemble 2-B' 2
  'Ensemble 2-B' 4
```

\* **\_atom\_sites\_alt\_gen.alt\_id**

This data item is a pointer to **\_atom\_sites\_alt.id** in the ATOM\_SITES\_ALT category.

\* **\_atom\_sites\_alt\_gen.ens\_id**

This data item is a pointer to **\_atom\_sites\_alt\_ens.id** in the ATOM\_SITES\_ALT\_ENS category.

**ATOM\_SITES\_FOOTNOTE**

Data items in the ATOM\_SITES\_FOOTNOTE category record detailed comments about an atom site or a group of atom sites.

Category group(s): **inclusive\_group**  
**atom\_group**

Category key(s): **\_atom\_sites\_footnote.id**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _atom_sites_footnote.id
  _atom_sites_footnote.text
  1
  ; The inhibitor binds to the enzyme in two alternative
  orientations. The two orientations have been assigned
  alternative IDs *1* and *2*.
  ;
  2
  ; Side chains of these residues adopt alternative
  orientations that correlate with the alternative
  orientations of the inhibitor.
  Side chains with alternative ID *1* and occupancy 0.58
  correlate with inhibitor orientation *1*.
  Side chains with alternative ID *2* and occupancy 0.42
  correlate with inhibitor orientation *2*.
  ;
  3
  ; The positions of these water molecules correlate with
  the alternative orientations of the inhibitor.
  Water molecules with alternative ID *1* and occupancy 0.58
  correlate with inhibitor orientation *1*.
  Water molecules with alternative ID *2* and occupancy 0.42
  correlate with inhibitor orientation *2*.
  ;
```

```
4
; Side chains of these residues adopt alternative
orientations that do not correlate with the alternative
orientation of the inhibitor.
;
5
; The positions of these water molecules correlate with
alternative orientations of amino-acid side chains that
do not correlate with alternative orientations of the
inhibitor.
;
```

\* **\_atom\_sites\_footnote.id** (code)

A code that identifies the footnote.

The following item(s) have an equivalent role in their respective categories:

**\_atom\_site.footnote\_id.**

Examples: 'a', 'b', '1', '2'. [atom\_sites\_footnote]

**\_atom\_sites\_footnote.text** (text)

The text of the footnote. Footnotes are used to describe an atom site or a group of atom sites in the ATOM\_SITE list. For example, footnotes may be used to indicate atoms for which the electron density is very weak, or atoms for which static disorder has been modelled.

[atom\_sites\_footnote]

**ATOM\_TYPE**

Data items in the ATOM\_TYPE category record details about the properties of the atoms that occupy the atom sites, such as the atomic scattering factors.

Category group(s): **inclusive\_group**  
**atom\_group**

Category key(s): **\_atom\_type.symbol**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _atom_type.symbol
  _atom_type.oxidation_number
  _atom_type.scat_Cromer_Mann_a1
  _atom_type.scat_Cromer_Mann_a2
  _atom_type.scat_Cromer_Mann_a3
  _atom_type.scat_Cromer_Mann_a4
  _atom_type.scat_Cromer_Mann_b1
  _atom_type.scat_Cromer_Mann_b2
  _atom_type.scat_Cromer_Mann_b3
  _atom_type.scat_Cromer_Mann_b4
  _atom_type.scat_Cromer_Mann_c
  C 0 2.31000 20.8439 1.02000 10.2075
  1.58860 0.568700 0.865000 51.6512 0.21560
  N 0 12.2126 0.005700 3.13220 9.89330
  2.01250 28.9975 1.16630 0.582600 -11.529
  O 0 3.04850 13.2771 2.28680 5.70110
  1.54630 0.323900 0.867000 32.9089 0.250800
  S 0 6.90530 1.46790 5.20340 22.2151
  1.43790 0.253600 1.58630 56.1720 0.866900
  CL -1 18.2915 0.006600 7.20840 1.17170
  6.53370 19.5424 2.33860 60.4486 -16.378
```

*Example 2 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].*

```
loop_
  _atom_type.symbol
  _atom_type.oxidation_number
  _atom_type.number_in_cell
  _atom_type.scat_dispersion_real
  _atom_type.scat_dispersion_imag
  _atom_type.scat_source
  C 0 72 .017 .009 International_Tables_Vol_IV_Table_2.2B
  H 0 100 0 0 International_Tables_Vol_IV_Table_2.2B
  O 0 12 .047 .032 International_Tables_Vol_IV_Table_2.2B
  N 0 4 .029 .018 International_Tables_Vol_IV_Table_2.2B
```

**\_atom\_type.analytical\_mass\_percent** (float)  
**\_atom\_type\_analytical\_mass\_%** (cif.core.dic 2.0.1)  
 Mass percentage of this atom type derived from chemical analysis.  
 The permitted range is [0, ∞). [atom\_type]

**\_atom\_type.description** (text)  
**\_atom\_type\_description** (cif.core.dic 2.0.1)  
 A description of the atom(s) designated by this atom type. In most cases, this is the element name and oxidation state of a single atom species. For disordered or nonstoichiometric structures it will describe a combination of atom species.  
 Examples: 'deuterium', '0.34Fe+0.66Ni'. [atom\_type]

**\_atom\_type.number\_in\_cell** (int)  
**\_atom\_type\_number\_in\_cell** (cif.core.dic 2.0.1)  
 Total number of atoms of this atom type in the unit cell.  
 The permitted range is [0, ∞). [atom\_type]

**\_atom\_type.oxidation\_number** (int)  
**\_atom\_type\_oxidation\_number** (cif.core.dic 2.0.1)  
 Formal oxidation state of this atom type in the structure.  
 The permitted range is [-8, 8]. Where no value is given, the assumed value is '0'. [atom\_type]

**\_atom\_type.radius\_bond** (float)  
**\_atom\_type\_radius\_bond** (cif.core.dic 2.0.1)  
 The effective intramolecular bonding radius in ångströms of this atom type.  
 The permitted range is [0.0, 5.0]. [atom\_type]

**\_atom\_type.radius\_contact** (float)  
**\_atom\_type\_radius\_contact** (cif.core.dic 2.0.1)  
 The effective intermolecular bonding radius in ångströms of this atom type.  
 The permitted range is [0.0, 5.0]. [atom\_type]

**\_atom\_type.scat\_Cromer\_Mann\_a1** (float)  
**\_atom\_type\_scatter\_Cromer\_Mann\_a1** (cif.core.dic 2.0.1)  
 The Cromer–Mann scattering-factor coefficient  $a_1$  used to calculate the scattering factors for this atom type.  
 References: *International Tables for X-ray Crystallography* (1974). Vol. IV, Table 2.2B; *International Tables for Crystallography* (2004). Vol. C, Tables 6.1.1.4 and 6.1.1.5. [atom\_type]

**\_atom\_type.scat\_Cromer\_Mann\_a2** (float)  
**\_atom\_type\_scatter\_Cromer\_Mann\_a2** (cif.core.dic 2.0.1)  
 The Cromer–Mann scattering-factor coefficient  $a_2$  used to calculate the scattering factors for this atom type.  
 References: *International Tables for X-ray Crystallography* (1974). Vol. IV, Table 2.2B; *International Tables for Crystallography* (2004). Vol. C, Tables 6.1.1.4 and 6.1.1.5. [atom\_type]

**\_atom\_type.scat\_Cromer\_Mann\_a3** (float)  
**\_atom\_type\_scatter\_Cromer\_Mann\_a3** (cif.core.dic 2.0.1)  
 The Cromer–Mann scattering-factor coefficient  $a_3$  used to calculate the scattering factors for this atom type.  
 References: *International Tables for X-ray Crystallography* (1974). Vol. IV, Table 2.2B; *International Tables for Crystallography* (2004). Vol. C, Tables 6.1.1.4 and 6.1.1.5. [atom\_type]

**\_atom\_type.scat\_Cromer\_Mann\_a4** (float)  
**\_atom\_type\_scatter\_Cromer\_Mann\_a4** (cif.core.dic 2.0.1)  
 The Cromer–Mann scattering-factor coefficient  $a_4$  used to calculate the scattering factors for this atom type.

References: *International Tables for X-ray Crystallography* (1974). Vol. IV, Table 2.2B; *International Tables for Crystallography* (2004). Vol. C, Tables 6.1.1.4 and 6.1.1.5. [atom\_type]

**\_atom\_type.scat\_Cromer\_Mann\_b1** (float)  
**\_atom\_type\_scatter\_Cromer\_Mann\_b1** (cif.core.dic 2.0.1)  
 The Cromer–Mann scattering-factor coefficient  $b_1$  used to calculate the scattering factors for this atom type.

References: *International Tables for X-ray Crystallography* (1974). Vol. IV, Table 2.2B; *International Tables for Crystallography* (2004). Vol. C, Tables 6.1.1.4 and 6.1.1.5. [atom\_type]

**\_atom\_type.scat\_Cromer\_Mann\_b2** (float)  
**\_atom\_type\_scatter\_Cromer\_Mann\_b2** (cif.core.dic 2.0.1)  
 The Cromer–Mann scattering-factor coefficient  $b_2$  used to calculate the scattering factors for this atom type.

References: *International Tables for X-ray Crystallography* (1974). Vol. IV, Table 2.2B; *International Tables for Crystallography* (2004). Vol. C, Tables 6.1.1.4 and 6.1.1.5. [atom\_type]

**\_atom\_type.scat\_Cromer\_Mann\_b3** (float)  
**\_atom\_type\_scatter\_Cromer\_Mann\_b3** (cif.core.dic 2.0.1)  
 The Cromer–Mann scattering-factor coefficient  $b_3$  used to calculate the scattering factors for this atom type.

References: *International Tables for X-ray Crystallography* (1974). Vol. IV, Table 2.2B; *International Tables for Crystallography* (2004). Vol. C, Tables 6.1.1.4 and 6.1.1.5. [atom\_type]

**\_atom\_type.scat\_Cromer\_Mann\_b4** (float)  
**\_atom\_type\_scatter\_Cromer\_Mann\_b4** (cif.core.dic 2.0.1)  
 The Cromer–Mann scattering-factor coefficient  $b_4$  used to calculate the scattering factors for this atom type.

References: *International Tables for X-ray Crystallography* (1974). Vol. IV, Table 2.2B; *International Tables for Crystallography* (2004). Vol. C, Tables 6.1.1.4 and 6.1.1.5. [atom\_type]

**\_atom\_type.scat\_Cromer\_Mann\_c** (float)  
**\_atom\_type\_scatter\_Cromer\_Mann\_c** (cif.core.dic 2.0.1)  
 The Cromer–Mann scattering-factor coefficient  $c$  used to calculate the scattering factors for this atom type.

References: *International Tables for X-ray Crystallography* (1974). Vol. IV, Table 2.2B; *International Tables for Crystallography* (2004). Vol. C, Tables 6.1.1.4 and 6.1.1.5. [atom\_type]

**\_atom\_type.scat\_dispersion\_imag** (float)  
**\_atom\_type\_scatter\_dispersion\_imag** (cif.core.dic 2.0.1)  
 The imaginary component of the anomalous-dispersion scattering factor,  $f''$ , in electrons for this atom type and the radiation identified by `_diffrn_radiation_wavelength.id`. [atom\_type]

**\_atom\_type.scat\_dispersion\_real** (float)  
**\_atom\_type\_scatter\_dispersion\_real** (cif.core.dic 2.0.1)  
 The real component of the anomalous-dispersion scattering factor,  $f'$ , in electrons for this atom type and the radiation identified by `_diffrn_radiation_wavelength.id`. [atom\_type]

**\_atom\_type.scat\_dispersion\_source** (text)  
 \_atom\_type\_scat\_dispersion\_source(cif\_core.dic 2.3)

Reference to the source of the real and imaginary dispersion corrections for scattering factors used for this atom type.

Example: 'International Tables Vol. IV Table 2.3.1'. [atom\_type]

**\_atom\_type.scat\_length\_neutron** (text)  
 \_atom\_type\_scat\_length\_neutron(cif\_core.dic 2.0.1)

The bound coherent scattering length in femtometres for the atom type at the isotopic composition used for the diffraction experiment.

[atom\_type]

**\_atom\_type.scat\_source** (text)  
 \_atom\_type\_scat\_source(cif\_core.dic 2.0.1)

Reference to the source of the scattering factors or scattering lengths used for this atom type.

Example: 'International Tables Vol. IV Table 2.4.6B'. [atom\_type]

**\_atom\_type.scat\_versus\_stol\_list** (text)  
 \_atom\_type\_scat\_versus\_stol\_list(cif\_core.dic 2.0.1)

A table of scattering factors as a function of  $(\sin \theta)/\lambda$ . This table should be well commented to indicate the items present. Regularly formatted lists are strongly recommended.

[atom\_type]

\* **\_atom\_type.symbol** (code)  
 \_atom\_type\_symbol(cif\_core.dic 2.0.1)

The code used to identify the atom species (singular or plural) representing this atom type. Normally this code is the element symbol. The code may be composed of any character except an underscore with the additional proviso that digits designate an oxidation state and must be followed by a + or - character.

The following item(s) have an equivalent role in their respective categories:

\_atom\_site.type\_symbol,  
 \_atom\_site\_anisotrop.type\_symbol,  
 \_chemical\_conn\_atom.type\_symbol,  
 \_chem\_comp\_atom.type\_symbol,  
 \_phasing\_MIR\_der\_site.atom\_type\_symbol.

Examples: 'C', 'Cu2+', 'H(SDS)', 'dummy', 'FeNi'. [atom\_type]

### AUDIT

Data items in the AUDIT category record details about the creation and subsequent updating of the data block. Note that these items apply only to the creation and updating of the data block, and should not be confused with the data items in the JOURNAL category that record different stages in the publication of the material in the data block.

Category group(s): inclusive\_group  
 audit\_group  
 Category key(s): \_audit.revision\_id

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
_audit.revision_id          1
_audit.creation_date        '1992-12-08'

_audit.creation_method
; Created by hand from PDB entry 5HVP, from the J. Biol.
  Chem. paper describing this structure and from
  laboratory records
;

_audit.update_record
; 1992-12-09  adjusted to reflect comments from B. McKeever
  1992-12-10  adjusted to reflect comments from H. Berman
  1992-12-12  adjusted to reflect comments from K. Watenpaugh
;
```

Example 2 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].

```
_audit.creation_date        1991-03-20
_audit.creation_method      from_xtal_archive_file_using_CIFIO
_audit.update_record
; 1991-04-09  text and data added by Tony Willis.
  1991-04-15  rec'd by co-editor as manuscript HL0007.
  1991-04-17  adjustments based on first referee report.
  1991-04-18  adjustments based on second referee report.
;
```

**\_audit.creation\_date** (yyyy-mm-dd)  
 \_audit\_creation\_date(cif\_core.dic 2.0.1)

A date that the data block was created. The date format is yyyy-mm-dd.

Example: '1990-07-12'. [audit]

**\_audit.creation\_method** (text)  
 \_audit\_creation\_method(cif\_core.dic 2.0.1)

A description of how data were entered into the data block.

Example: 'spawned by the program QBEE'. [audit]

\* **\_audit.revision\_id** (code)

The value of \_audit.revision\_id must uniquely identify a record in the AUDIT list.

Example: 'rev1'. [audit]

**\_audit.update\_record** (text)  
 \_audit\_update\_record(cif\_core.dic 2.0.1)

A record of any changes to the data block. The update format is a date (yyyy-mm-dd) followed by a description of the changes. The latest update entry is added to the bottom of this record.

Example: '1990-07-15 Updated by the Co-editor'. [audit]

### AUDIT\_AUTHOR

Data items in the AUDIT\_AUTHOR category record details about the author(s) of the data block.

Category group(s): inclusive\_group  
 audit\_group  
 Category key(s): \_audit\_author.name

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
loop_
_audit_author.name
_audit_author.address
'Fitzgerald, Paula M.D.'
; Department of Biophysical Chemistry
  Merck Research Laboratories
  P. O. Box 2000, Ry80M203
  Rahway, New Jersey 07065
  USA
;
'McKeever, Brian M.'
; Department of Biophysical Chemistry
  Merck Research Laboratories
  P. O. Box 2000, Ry80M203
  Rahway, New Jersey 07065
  USA
;
'Van Middlesworth, J.F.'
; Department of Biophysical Chemistry
  Merck Research Laboratories
  P. O. Box 2000, Ry80M203
  Rahway, New Jersey 07065
  USA
;
'Springer, James P.'
; Department of Biophysical Chemistry
  Merck Research Laboratories
  P. O. Box 2000, Ry80M203
  Rahway, New Jersey 07065
  USA
;
```

**\_audit\_author.address** (text)  
**\_audit\_author\_address** (cif\_core.dic 2.0.1)

The address of an author of this data block. If there are multiple authors, **\_audit\_author.address** is looped with **\_audit\_author.name**.

Example:

```
; Department
  Institute
  Street
  City and postcode
  COUNTRY
; [audit_author]
```

\***\_audit\_author.name** (line)  
**\_audit\_author\_name** (cif\_core.dic 2.0.1)

The name of an author of this data block. If there are multiple authors, **\_audit\_author.name** is looped with **\_audit\_author.address**. The family name(s), followed by a comma and including any dynastic components, precedes the first name(s) or initial(s).

Examples: 'Bleary, Percival R.', 'O'Neil, F.K.', 'Van den Bossche, G.', 'Yang, D.-L.', 'Simonov, Yu.A.'. [audit\_author]

### AUDIT\_CONFORM

Data items in the AUDIT\_CONFORM category describe the dictionary versions against which the data names appearing in the current data block are conformant.

Category group(s): **inclusive\_group**  
**audit\_group**  
 Category key(s): **\_audit\_conform.dict\_name**  
**\_audit\_conform.dict\_version**

Example 1 – any file conforming to the current CIF core dictionary.

```
_audit_conform.dict_name      cif_core.dic
_audit_conform.dict_version    2.3.1
_audit_conform.dict_location
  ftp://ftp.iucr.org/pub/cif_core.2.3.1.dic
```

**\_audit\_conform.dict\_location** (text)  
**\_audit\_conform\_dict\_location** (cif\_core.dic 2.0.1)

A file name or uniform resource locator (URL) for the dictionary to which the current data block conforms.

[audit\_conform]

**\_audit\_conform.dict\_name** (text)  
**\_audit\_conform\_dict\_name** (cif\_core.dic 2.0.1)

The string identifying the highest-level dictionary defining data names used in this file.

[audit\_conform]

**\_audit\_conform.dict\_version** (text)  
**\_audit\_conform\_dict\_version** (cif\_core.dic 2.0.1)

The version number of the dictionary to which the current data block conforms.

[audit\_conform]

### AUDIT\_CONTACT\_AUTHOR

Data items in the AUDIT\_CONTACT\_AUTHOR category record details about the name and address of the author to be contacted concerning the content of this data block.

Category group(s): **inclusive\_group**  
**audit\_group**  
 Category key(s): **\_audit\_contact\_author.name**

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
_audit_contact_author.name    'Fitzgerald, Paula M.D.'
_audit_contact_author.address
; Department of Biophysical Chemistry
  Merck Research Laboratories
  PO Box 2000, Ry80M203
  Rahway, New Jersey 07065
  USA
;
_audit_contact_author.phone    '1(908)5945510'
_audit_contact_author.fax      '1(908)5946645'
_audit_contact_author.email    'paula_fitzgerald@merck.com'
```

**\_audit\_contact\_author.address** (text)  
**\_audit\_contact\_author\_address** (cif\_core.dic 2.0.1)

The mailing address of the author of the data block to whom correspondence should be addressed.

Example:

```
; Department
  Institute
  Street
  City and postcode
  COUNTRY
; [audit_contact_author]
```

**\_audit\_contact\_author.email** (line)  
**\_audit\_contact\_author\_email** (cif\_core.dic 2.0.1)

The electronic mail address of the author of the data block to whom correspondence should be addressed, in a form recognizable to international networks. The format of e-mail addresses is given in Section 3.4, *Address Specification*, of *Internet Message Format*, RFC 2822, P. Resnick (Editor), Network Standards Group, April 2001.

Examples: 'name@host.domain.country', 'bm@iucr.org'.

[audit\_contact\_author]

**\_audit\_contact\_author.fax** (line)  
**\_audit\_contact\_author\_fax** (cif\_core.dic 2.0.1)

The facsimile telephone number of the author of the data block to whom correspondence should be addressed. The recommended style starts with the international dialing prefix, followed by the area code in parentheses, followed by the local number with no spaces.

Examples: '12(34)9477334', '12()349477334'. [audit\_contact\_author]

\***\_audit\_contact\_author.name** (line)  
**\_audit\_contact\_author\_name** (cif\_core.dic 2.0.1)

The name of the author of the data block to whom correspondence should be addressed. The family name(s), followed by a comma and including any dynastic components, precedes the first name(s) or initial(s).

Examples: 'Bleary, Percival R.', 'O'Neil, F.K.', 'Van den Bossche, G.', 'Yang, D.-L.', 'Simonov, Yu.A.'. [audit\_contact\_author]

**\_audit\_contact\_author.phone** (line)  
**\_audit\_contact\_author\_phone** (cif\_core.dic 2.0.1)

The telephone number of the author of the data block to whom correspondence should be addressed. The recommended style starts with the international dialing prefix, followed by the area code in parentheses, followed by the local number and any extension number prefixed by 'x', with no spaces.

Examples: '12(34)9477330', '12()349477330', '12(34)9477330x5543'.

[audit\_contact\_author]

**AUDIT\_LINK**

Data items in the AUDIT\_LINK category record details about the relationships between data blocks in the current CIF.

Category key(s): `_audit_link.block_code`  
`_audit_link.block_description`

*Example 1 – multiple structure paper, as illustrated in A Guide to CIF for Authors (1995). IUCr: Chester.*

```
loop_
_audit_link.block_code
_audit_link.block_description
.      'discursive text of paper with two structures'
morA_ (1) 'structure 1 of 2'
morA_ (2) 'structure 2 of 2'
```

*Example 2 – example file for the one-dimensional incommensurately modulated structure of K<sub>2</sub>SeO<sub>4</sub>.*

```
loop_
_audit_link.block_code
_audit_link.block_description
.      'publication details'
KSE_COM 'experimental data common to ref./mod. structures'
KSE_REF 'reference structure'
KSE_MOD 'modulated structure'
```

\* `_audit_link.block_code` (code)

`_audit_link_block_code` (cif.core.dic 2.3)

The value of `_audit_block.code` associated with a data block in the current file related to the current data block. The special value '.' may be used to refer to the current data block for completeness.

[audit\_link]

\* `_audit_link.block_description` (text)

`_audit_link_block_description` (cif.core.dic 2.3)

A textual description of the relationship of the referenced data block to the current one.

[audit\_link]

**CELL**

Data items in the CELL category record details about the crystallographic cell parameters.

Category group(s): `inclusive_group`  
`cell_group`

Category key(s): `_cell.entry_id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
_cell.entry_id      '5HVP'
_cell.length_a      58.39
_cell.length_a_esd  0.05
_cell.length_b      86.70
_cell.length_b_esd  0.12
_cell.length_c      46.27
_cell.length_c_esd  0.06
_cell.angle_alpha    90.00
_cell.angle_beta     90.00
_cell.angle_gamma    90.00
_cell.volume         234237
```

`_cell.details`  
; The cell parameters were refined every twenty frames during data integration. The cell lengths given are the mean of 55 such refinements; the esds given are the root mean square deviations of these 55 observations from that mean.

;

*Example 2 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].*

```
_cell.length_a      5.959
_cell.length_a_esd  0.001
_cell.length_b      14.956
_cell.length_b_esd  0.001
_cell.length_c      19.737
_cell.length_c_esd  0.003
_cell.angle_alpha    90.0
_cell.angle_beta     90.0
_cell.angle_gamma    90.0
_cell.volume         1759.0
_cell.volume_esd     0.3
```

`_cell.angle_alpha` (float, su)  
`_cell_angle_alpha` (cif.core.dic 2.0.1)

Unit-cell angle  $\alpha$  of the reported structure in degrees.

The permitted range is [0.0, 180.0].

Related item: `_cell.angle_alpha_esd` (associated esd). Where no value is given, the assumed value is '90.0'. [cell]

`_cell.angle_alpha_esd` (float)

The standard uncertainty (estimated standard deviation) of `_cell.angle_alpha`.

Related item: `_cell.angle_alpha` (associated value). [cell]

`_cell.angle_beta` (float, su)

`_cell_angle_beta` (cif.core.dic 2.0.1)

Unit-cell angle  $\beta$  of the reported structure in degrees.

The permitted range is [0.0, 180.0].

Related item: `_cell.angle_beta_esd` (associated esd). Where no value is given, the assumed value is '90.0'. [cell]

`_cell.angle_beta_esd` (float)

The standard uncertainty (estimated standard deviation) of `_cell.angle_beta`.

Related item: `_cell.angle_beta` (associated value). [cell]

`_cell.angle_gamma` (float, su)

`_cell_angle_gamma` (cif.core.dic 2.0.1)

Unit-cell angle  $\gamma$  of the reported structure in degrees.

The permitted range is [0.0, 180.0].

Related item: `_cell.angle_gamma_esd` (associated esd). Where no value is given, the assumed value is '90.0'. [cell]

`_cell.angle_gamma_esd` (float)

The standard uncertainty (estimated standard deviation) of `_cell.angle_gamma`.

Related item: `_cell.angle_gamma` (associated value). [cell]

\* `_cell.details` (text)

`_cell_special_details` (cif.core.dic 2.0.1)

A description of special aspects of the cell choice, noting possible alternative settings.

Examples: 'pseudo-orthorhombic',  
'standard setting from 45 deg rotation around c'. [cell]

\* `_cell.entry_id`

This data item is a pointer to `_entry.id` in the ENTRY category.

`_cell.formula_units_Z` (int)

`_cell_formula_units_Z` (cif.core.dic 2.0.1)

The number of the formula units in the unit cell as specified by `_chemical_formula.structural`, `_chemical_formula.moiety` or `_chemical_formula.sum`.

The permitted range is [1,  $\infty$ ). [cell]

**\_cell.length\_a** (float, su)  
 \_cell\_length\_a (cif\_core.dic 2.0.1)

Unit-cell length *a* corresponding to the structure reported in ångströms.

The permitted range is [0.0, ∞).

Related item: **\_cell.length\_a\_esd** (associated esd). [cell]

**\_cell.length\_a\_esd** (float)  
 The standard uncertainty (estimated standard deviation) of **\_cell.length\_a**.

Related item: **\_cell.length\_a** (associated value). [cell]

**\_cell.length\_b** (float, su)  
 \_cell\_length\_b (cif\_core.dic 2.0.1)

Unit-cell length *b* corresponding to the structure reported in ångströms.

The permitted range is [0.0, ∞).

Related item: **\_cell.length\_b\_esd** (associated esd). [cell]

**\_cell.length\_b\_esd** (float)  
 The standard uncertainty (estimated standard deviation) of **\_cell.length\_b**.

Related item: **\_cell.length\_b** (associated value). [cell]

**\_cell.length\_c** (float, su)  
 \_cell\_length\_c (cif\_core.dic 2.0.1)

Unit-cell length *c* corresponding to the structure reported in ångströms.

The permitted range is [0.0, ∞).

Related item: **\_cell.length\_c\_esd** (associated esd). [cell]

**\_cell.length\_c\_esd** (float)  
 The standard uncertainty (estimated standard deviation) of **\_cell.length\_c**.

Related item: **\_cell.length\_c** (associated value). [cell]

**\_cell.reciprocal\_angle\_alpha** (float, su)  
 \_cell\_reciprocal\_angle\_alpha (cif\_core.dic 2.3)

The angle  $\alpha^*$  defining the reciprocal cell in degrees.  $\alpha^*$ ,  $\beta^*$  and  $\gamma^*$  are related to the angles in the real cell by

$$\begin{aligned}\cos \alpha^* &= (\cos \beta \cos \gamma - \cos \alpha) / (\sin \beta \sin \gamma), \\ \cos \beta^* &= (\cos \gamma \cos \alpha - \cos \beta) / (\sin \gamma \sin \alpha), \\ \cos \gamma^* &= (\cos \alpha \cos \beta - \cos \gamma) / (\sin \alpha \sin \beta).\end{aligned}$$

Reference: Buerger, M. J. (1942). *X-ray Crystallography*, p. 360. New York: John Wiley & Sons Inc.

The permitted range is [0.0, 180.0].

Related item: **\_cell.reciprocal\_angle\_alpha\_esd** (associated esd). Where no value is given, the assumed value is '90.0'. [cell]

**\_cell.reciprocal\_angle\_alpha\_esd** (float)  
 The estimated standard deviation of **\_cell.reciprocal\_angle\_alpha**.

Related item: **\_cell.reciprocal\_angle\_alpha** (associated value). [cell]

**\_cell.reciprocal\_angle\_beta** (float, su)  
 \_cell\_reciprocal\_angle\_beta (cif\_core.dic 2.3)

The angle  $\beta^*$  defining the reciprocal cell in degrees.  $\alpha^*$ ,  $\beta^*$  and  $\gamma^*$  are related to the angles in the real cell by

$$\begin{aligned}\cos \alpha^* &= (\cos \beta \cos \gamma - \cos \alpha) / (\sin \beta \sin \gamma), \\ \cos \beta^* &= (\cos \gamma \cos \alpha - \cos \beta) / (\sin \gamma \sin \alpha), \\ \cos \gamma^* &= (\cos \alpha \cos \beta - \cos \gamma) / (\sin \alpha \sin \beta).\end{aligned}$$

Reference: Buerger, M. J. (1942). *X-ray Crystallography*, p. 360. New York: John Wiley & Sons Inc.

The permitted range is [0.0, 180.0].

Related item: **\_cell.reciprocal\_angle\_beta\_esd** (associated esd). Where no value is given, the assumed value is '90.0'. [cell]

**\_cell.reciprocal\_angle\_beta\_esd** (float)

The estimated standard deviation of **\_cell.reciprocal\_angle\_beta**.

Related item: **\_cell.reciprocal\_angle\_beta** (associated value). [cell]

**\_cell.reciprocal\_angle\_gamma** (float, su)  
 \_cell\_reciprocal\_angle\_gamma (cif\_core.dic 2.3)

The angle  $\gamma^*$  defining the reciprocal cell in degrees.  $\alpha^*$ ,  $\beta^*$  and  $\gamma^*$  are related to the angles in the real cell by

$$\begin{aligned}\cos \alpha^* &= (\cos \beta \cos \gamma - \cos \alpha) / (\sin \beta \sin \gamma), \\ \cos \beta^* &= (\cos \gamma \cos \alpha - \cos \beta) / (\sin \gamma \sin \alpha), \\ \cos \gamma^* &= (\cos \alpha \cos \beta - \cos \gamma) / (\sin \alpha \sin \beta).\end{aligned}$$

Reference: Buerger, M. J. (1942). *X-ray Crystallography*, p. 360. New York: John Wiley & Sons Inc.

The permitted range is [0.0, 180.0].

Related item: **\_cell.reciprocal\_angle\_gamma\_esd** (associated esd). Where no value is given, the assumed value is '90.0'. [cell]

**\_cell.reciprocal\_angle\_gamma\_esd** (float)

The estimated standard deviation of **\_cell.reciprocal\_angle\_gamma**.

Related item: **\_cell.reciprocal\_angle\_gamma** (associated value). [cell]

**\_cell.reciprocal\_length\_a** (float, su)  
 \_cell\_reciprocal\_length\_a (cif\_core.dic 2.3)

The reciprocal-cell length *a\** in inverse ångströms. *a\**, *b\** and *c\** are related to the lengths in the real cell by

$$\begin{aligned}a^* &= bc \sin \alpha / V, \\ b^* &= ca \sin \beta / V, \\ c^* &= ab \sin \gamma / V,\end{aligned}$$

where *V* is the cell volume.

Reference: Buerger, M. J. (1942). *X-ray Crystallography*, p. 360. New York: John Wiley & Sons Inc.

The permitted range is [0.0, ∞).

Related item: **\_cell.reciprocal\_length\_a\_esd** (associated esd). [cell]

**\_cell.reciprocal\_length\_a\_esd** (float)

The estimated standard deviation of **\_cell.reciprocal\_length\_a**.

Related item: **\_cell.reciprocal\_length\_a** (associated value). [cell]

**`_cell.reciprocal_length_b`** (float, su)  
`_cell_reciprocal_length_b` (cif\_core.dic 2.3)

The reciprocal-cell length  $b^*$  in inverse ångströms.  $a^*$ ,  $b^*$  and  $c^*$  are related to the lengths in the real cell by

$$\begin{aligned} a^* &= bc \sin \alpha / V, \\ b^* &= ca \sin \beta / V, \\ c^* &= ab \sin \gamma / V, \end{aligned}$$

where  $V$  is the cell volume.

Reference: Buerger, M. J. (1942). *X-ray Crystallography*, p. 360. New York: John Wiley & Sons Inc.

The permitted range is [0.0, ∞).

Related item: `_cell.reciprocal_length_b_esd` (associated esd). [cell]

**`_cell.reciprocal_length_b_esd`** (float)

The estimated standard deviation of `_cell.reciprocal_length_b`.

Related item: `_cell.reciprocal_length_b` (associated value). [cell]

**`_cell.reciprocal_length_c`** (float, su)

`_cell_reciprocal_length_c` (cif\_core.dic 2.3)

The reciprocal-cell length  $c^*$  in inverse ångströms.  $a^*$ ,  $b^*$  and  $c^*$  are related to the lengths in the real cell by

$$\begin{aligned} a^* &= bc \sin \alpha / V, \\ b^* &= ca \sin \beta / V, \\ c^* &= ab \sin \gamma / V, \end{aligned}$$

where  $V$  is the cell volume.

Reference: Buerger, M. J. (1942). *X-ray Crystallography*, p. 360. New York: John Wiley & Sons Inc.

The permitted range is [0.0, ∞).

Related item: `_cell.reciprocal_length_c_esd` (associated esd). [cell]

**`_cell.reciprocal_length_c_esd`** (float)

The estimated standard deviation of `_cell.reciprocal_length_c`.

Related item: `_cell.reciprocal_length_c` (associated value). [cell]

**`_cell.volume`** (float, su)

`_cell_volume` (cif\_core.dic 2.0.1)

Cell volume  $V$  in ångströms cubed.

$$V = abc(1 - \cos^2 \alpha - \cos^2 \beta - \cos^2 \gamma + 2 \cos \alpha \cos \beta \cos \gamma)^{1/2},$$

where  $a = \text{\_cell.length\_a}$ ,  $b = \text{\_cell.length\_b}$ ,  $c = \text{\_cell.length\_c}$ ,  $\alpha = \text{\_cell.angle\_alpha}$ ,  $\beta = \text{\_cell.angle\_beta}$  and  $\gamma = \text{\_cell.angle\_gamma}$ .

The permitted range is [0.0, ∞).

Related item: `_cell.volume_esd` (associated esd). [cell]

**`_cell.volume_esd`** (float)

The standard uncertainty (estimated standard deviation) of `_cell.volume`.

Related item: `_cell.volume` (associated value). [cell]

**`_cell.Z_PDB`** (int)

The number of the polymeric chains in a unit cell. In the case of heteropolymers,  $Z$  is the number of occurrences of the most populous chain. This data item is provided for compatibility with the original Protein Data Bank format, and only for that purpose.

The permitted range is [1, ∞). [cell]

## CELL\_MEASUREMENT

Data items in the CELL\_MEASUREMENT category record details about the measurement of the crystallographic cell parameters.

Category group(s): `inclusive_group`

`cell_group`

Category key(s): `_cell_measurement.entry_id`

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

<code>_cell_measurement.entry_id</code>	'5HVP'
<code>_cell_measurement.temp</code>	293
<code>_cell_measurement.temp_esd</code>	3
<code>_cell_measurement.theta_min</code>	11
<code>_cell_measurement.theta_max</code>	31
<code>_cell_measurement.wavelength</code>	1.54

Example 2 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].

<code>_cell_measurement.temp</code>	293
<code>_cell_measurement.reflns_used</code>	25
<code>_cell_measurement.theta_min</code>	25
<code>_cell_measurement.theta_max</code>	31

\* `_cell_measurement.entry_id`

This data item is a pointer to `_entry.id` in the ENTRY category.

**`_cell_measurement.pressure`** (float, su)

`_cell_measurement_pressure` (cif\_core.dic 2.0.1)

The pressure in kilopascals at which the unit-cell parameters were measured (not the pressure at which the sample was synthesized).

Related item: `_cell_measurement.pressure_esd` (associated esd).

[cell\_measurement]

**`_cell_measurement.pressure_esd`** (float)

The standard uncertainty (estimated standard deviation) of `_cell_measurement.pressure`.

Related item: `_cell_measurement.pressure` (associated value).

[cell\_measurement]

**`_cell_measurement.radiation`** (line)

`_cell_measurement_radiation` (cif\_core.dic 2.0.1)

Description of the radiation used to measure the unit-cell data. See also `_cell_measurement.wavelength`.

Examples: 'neutron', 'Cu K\alpha', 'synchrotron'. [cell\_measurement]

**`_cell_measurement.reflns_used`** (int)

`_cell_measurement_reflns_used` (cif\_core.dic 2.0.1)

The total number of reflections used to determine the unit cell. These reflections may be specified as CELL\_MEASUREMENT\_REFLN data items.

[cell\_measurement]

**`_cell_measurement.temp`** (float, su)

`_cell_measurement_temperature` (cif\_core.dic 2.0.1)

The temperature in kelvins at which the unit-cell parameters were measured (not the temperature of synthesis).

The permitted range is [0.0, ∞).

Related item: `_cell_measurement.temp_esd` (associated esd).

[cell\_measurement]

**`_cell_measurement.temp_esd`** (float)

The standard uncertainty (estimated standard deviation) of `_cell_measurement.temp`.

Related item: `_cell_measurement.temp` (associated value). [cell\_measurement]

**\_cell\_measurement.theta\_max** (float)  
 \_cell\_measurement\_theta\_max (cif\_core.dic 2.0.1)

The maximum  $\theta$  angle of reflections used to measure the unit cell in degrees.

The permitted range is [0.0, 90.0]. [cell\_measurement]

**\_cell\_measurement.theta\_min** (float)  
 \_cell\_measurement\_theta\_min (cif\_core.dic 2.0.1)

The minimum  $\theta$  angle of reflections used to measure the unit cell in degrees.

The permitted range is [0.0, 90.0]. [cell\_measurement]

**\_cell\_measurement.wavelength** (float)  
 \_cell\_measurement\_wavelength (cif\_core.dic 2.0.1)

The wavelength in ångströms of the radiation used to measure the unit cell. If this is not specified, the wavelength is assumed to be that specified in the category DIFFRN\_RADIATION\_WAVELENGTH.

The permitted range is [0.0,  $\infty$ ]. [cell\_measurement]

### CELL\_MEASUREMENT\_REFLN

Data items in the CELL\_MEASUREMENT\_REFLN category record details about the reflections used to determine the crystallographic cell parameters. The CELL\_MEASUREMENT\_REFLN data items would in general be used only for diffractometer data.

Category group(s): inclusive\_group  
 cell\_group  
 Category key(s): \_cell\_measurement\_refl\_index\_h  
 \_cell\_measurement\_refl\_index\_k  
 \_cell\_measurement\_refl\_index\_l

Example 1 – extracted from the CAD-4 listing of  $Rb_2S_2O_6$  at room temperature (unpublished).

```
loop_
_cell_measurement_refl_index_h
_cell_measurement_refl_index_k
_cell_measurement_refl_index_l
_cell_measurement_refl_theta
  -2  4  1  8.67
   0  3  2  9.45
   3  0  2  9.46
  -3  4  1  8.93
  -2  1  -2  7.53
  10  0  0  23.77
   0 10  0  23.78
  -5  4  1  11.14
# - - - data truncated for brevity - - -
```

\* **\_cell\_measurement\_refl\_index\_h** (int)  
 \_cell\_measurement\_refl\_index\_h (cif\_core.dic 2.0.1)

Miller index  $h$  of a reflection used for measurement of the unit cell.  
 [cell\_measurement\_refl]

\* **\_cell\_measurement\_refl\_index\_k** (int)  
 \_cell\_measurement\_refl\_index\_k (cif\_core.dic 2.0.1)

Miller index  $k$  of a reflection used for measurement of the unit cell.  
 [cell\_measurement\_refl]

\* **\_cell\_measurement\_refl\_index\_l** (int)  
 \_cell\_measurement\_refl\_index\_l (cif\_core.dic 2.0.1)

Miller index  $l$  of a reflection used for measurement of the unit cell.  
 [cell\_measurement\_refl]

**\_cell\_measurement\_refl.theta** (float)  
 \_cell\_measurement\_refl\_theta (cif\_core.dic 2.0.1)

$\theta$  angle for a reflection used for measurement of the unit cell in degrees.

The permitted range is [0.0, 90.0]. [cell\_measurement\_refl]

### CHEM\_COMP

Data items in the CHEM\_COMP category give details about each of the chemical components from which the relevant chemical structures can be constructed, such as name, mass or charge. The related categories CHEM\_COMP\_ATOM, CHEM\_COMP\_BOND, CHEM\_COMP\_ANGLE etc. describe the detailed geometry of these chemical components.

Category group(s): inclusive\_group  
 chem\_comp\_group  
 Category key(s): \_chem\_comp.id

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
loop_
_chem_comp.id
_chem_comp.model_source
_chem_comp.name
  phe '1987 Protin/Prolsq Ideals file' phenylalanine
  val '1987 Protin/Prolsq Ideals file' alanine
# - - - data truncated for brevity - - -
```

**\_chem\_comp.formula** (text)

The formula for the chemical component. Formulae are written according to the following rules: (1) Only recognized element symbols may be used. (2) Each element symbol is followed by a 'count' number. A count of '1' may be omitted. (3) A space or parenthesis must separate each cluster of (element symbol + count), but in general parentheses are not used. (4) The order of elements depends on whether carbon is present or not. If carbon is present, the order should be: C, then H, then the other elements in alphabetical order of their symbol. If carbon is not present, the elements are listed purely in alphabetic order of their symbol. This is the 'Hill' system used by *Chemical Abstracts*.

Example: 'C18 H19 N7 O8 S'. [chem\_comp]

**\_chem\_comp.formula\_weight** (float)

Formula mass in daltons of the chemical component.

The permitted range is [1.0,  $\infty$ ]. [chem\_comp]

\* **\_chem\_comp.id** (ucode)

The value of **\_chem\_comp.id** must uniquely identify each item in the CHEM\_COMP list. For protein polymer entities, this is the three-letter code for the amino acid. For nucleic acid polymer entities, this is the one-letter code for the base.

The following item(s) have an equivalent role in their respective categories:

```
_atom_site.label_comp_id,
_chem_comp.mon_nstd_parent_comp_id,
_chem_comp_atom.comp_id,
_chem_comp_chir.comp_id,
_chem_comp_chir_atom.comp_id,
_chem_comp_plane.comp_id,
_chem_comp_plane_atom.comp_id,
_entity_poly_seq.mon_id,
_chem_comp_angle.comp_id,
_chem_comp_bond.comp_id,
_chem_comp_tor.comp_id,
_chem_comp_tor_value.comp_id,
_geom_angle.atom_site_label_comp_id_1,
_geom_angle.atom_site_label_comp_id_2,
_geom_angle.atom_site_label_comp_id_3,
_geom_bond.atom_site_label_comp_id_1,
_geom_bond.atom_site_label_comp_id_2,
_geom_contact.atom_site_label_comp_id_1,
_geom_contact.atom_site_label_comp_id_2,
_geom_hbond.atom_site_label_comp_id_A,
_geom_hbond.atom_site_label_comp_id_D,
```

`_geom_hbond.atom_site_label_comp_id_H`,  
`_geom_torsion.atom_site_label_comp_id_1`,  
`_geom_torsion.atom_site_label_comp_id_2`,  
`_geom_torsion.atom_site_label_comp_id_3`,  
`_geom_torsion.atom_site_label_comp_id_4`,  
`_struct_conf.beg_label_comp_id`,  
`_struct_conf.end_label_comp_id`,  
`_struct_conn.ptnr1_label_comp_id`,  
`_struct_conn.ptnr2_label_comp_id`,  
`_struct_mon_nucl.label_comp_id`,  
`_struct_mon_prot.label_comp_id`,  
`_struct_mon_prot_cis.label_comp_id`,  
`_struct_ncs_dom_lim.beg_label_comp_id`,  
`_struct_ncs_dom_lim.end_label_comp_id`,  
`_struct_ref_seq_dif.db_mon_id`,  
`_struct_ref_seq_dif.mon_id`,  
`_struct_sheet_range.beg_label_comp_id`,  
`_struct_sheet_range.end_label_comp_id`,  
`_struct_site_gen.label_comp_id`.

Examples: 'ala', 'val', 'A', 'C'.

[chem\_comp]

**`_chem_comp.model_details`** (text)

A description of special aspects of the generation of the coordinates for the model of the component.

Example: 'geometry idealized but not minimized'.

[chem\_comp]

**`_chem_comp.model_eref`** (line)

A pointer to an external reference file from which the atomic description of the component is taken.

[chem\_comp]

**`_chem_comp.model_source`** (text)

The source of the coordinates for the model of the component.

Examples: 'CSD entry ABCDEF', 'built using Quanta/Charmm'.

[chem\_comp]

**`_chem_comp.mon_nstd_class`** (text)

A description of the class of a nonstandard monomer if the nonstandard monomer represents a modification of a standard monomer.

Examples: 'iodinated base', 'phosphorylated amino acid', 'brominated base', 'modified amino acid', 'glycosylated amino acid'.

[chem\_comp]

**`_chem_comp.mon_nstd_details`** (text)

A description of special details of a nonstandard monomer.

[chem\_comp]

**`_chem_comp.mon_nstd_flag`** (ucode)

'yes' indicates that this is a 'standard' monomer, 'no' indicates that it is 'nonstandard'. Nonstandard monomers should be described in more detail using the `_chem_comp.mon_nstd_parent`, `_chem_comp.mon_nstd_class` and `_chem_comp.mon_nstd_details` data items.

The data value must be one of the following:

no	the monomer is nonstandard
n	abbreviation for 'no'
yes	the monomer is standard
y	abbreviation for 'yes'

Where no value is given, the assumed value is 'no'.

[chem\_comp]

**`_chem_comp.mon_nstd_parent`** (code)

The name of the parent monomer of the nonstandard monomer, if the nonstandard monomer represents a modification of a standard monomer.

Examples: 'tyrosine', 'cytosine'.

[chem\_comp]

**`_chem_comp.mon_nstd_parent_comp_id`**

The identifier for the parent component of the nonstandard component. This data item is a pointer to `_chem_comp.id` in the CHEM\_COMP category.

**`_chem_comp.name`** (line)

The full name of the component.

Examples: 'alanine', 'valine', 'adenine', 'cytosine'.

[chem\_comp]

**`_chem_comp.number_atoms_all`** (int)

The total number of atoms in the component.

The permitted range is [1, ∞).

[chem\_comp]

**`_chem_comp.number_atoms_nh`** (int)

The number of non-hydrogen atoms in the component.

The permitted range is [1, ∞).

[chem\_comp]

**`_chem_comp.one_letter_code`** (uchar1)

For standard polymer components, the one-letter code for the component. If there is not a standard one-letter code for this component, or if this is a non-polymer component, the one-letter code should be given as 'X'. This code may be preceded by a '+' character to indicate that the component is a modification of a standard component.

Examples: 'A' (alanine or adenine), 'B' (ambiguous asparagine/aspartic acid), 'R' (arginine), 'N' (asparagine), 'D' (aspartic acid), 'C' (cysteine or cystine or cytosine), 'Q' (glutamine), 'E' (glutamic acid), 'Z' (ambiguous glutamine/glutamic acid), 'G' (glycine or guanine), 'H' (histidine), 'I' (isoleucine), 'L' (leucine), 'K' (lysine), 'M' (methionine), 'F' (phenylalanine), 'P' (proline), 'S' (serine), 'T' (threonine or thymine), 'W' (tryptophan), 'Y' (tyrosine), 'V' (valine), 'U' (uracil), 'O' (water), 'X' (other).

[chem\_comp]

**`_chem_comp.three_letter_code`** (uchar3)

For standard polymer components, the three-letter code for the component. If there is not a standard three-letter code for this component, or if this is a non-polymer component, the three-letter code should be given as 'UNK'. This code may be preceded by a '+' character to indicate that the component is a modification of a standard component.

Examples: 'ALA' (alanine), 'ARG' (arginine), 'ASN' (asparagine), 'ASP' (aspartic acid), 'ASX' (ambiguous asparagine/aspartic acid), 'CYS' (cysteine), 'GLN' (glutamine), 'GLU' (glutamic acid), 'GLY' (glycine), 'GLX' (ambiguous glutamine/glutamic acid), 'HIS' (histidine), 'ILE' (isoleucine), 'LEU' (leucine), 'LYS' (lysine), 'MET' (methionine), 'PHE' (phenylalanine), 'PRO' (proline), 'SER' (serine), 'THR' (threonine), 'TRP' (tryptophan), 'TRY' (tyrosine), 'VAL' (valine), '1MA' (1-methyladenosine), '5MC' (5-methylcytosine), 'OMC' (2'-O-methylcytosine), '1MG' (1-methylguanosine), '2MG' (N(2)-methylguanosine), 'M2G' (N(2)-dimethylguanosine), '7MG' (7-methylguanosine), 'OMG' (2'-O-methylguanosine), 'H2U' (dihydrouridine), '5MU' (ribosylthymidine), 'PSU' (pseudouridine), 'ACE' (acetic acid), 'FOR' (formic acid), 'HOH' (water), 'UNK' (other).

[chem\_comp]

\* **\_chem\_comp.type** (uline)

For standard polymer components, the type of the monomer. Note that monomers that will form polymers are of three types: linking monomers, monomers with some type of N-terminal (or 5') cap and monomers with some type of C-terminal (or 3') cap.

The following item(s) have an equivalent role in their respective categories:

\_chem\_comp.link.type\_comp\_1  
\_chem\_comp.link.type\_comp\_2

The data value must be one of the following:

```
'D-peptide linking'
'L-peptide linking'
'D-peptide NH3 amino terminus'
'L-peptide NH3 amino terminus'
'D-peptide COOH carboxy terminus'
'L-peptide COOH carboxy terminus'
'DNA linking'
'RNA linking'
'DNA OH 5 prime terminus'
'RNA OH 5 prime terminus'
'DNA OH 3 prime terminus'
'RNA OH 3 prime terminus'
'D-saccharide 1,4 and 1,4 linking'
'L-saccharide 1,4 and 1,4 linking'
'D-saccharide 1,4 and 1,6 linking'
'L-saccharide 1,4 and 1,6 linking'
L-saccharide
D-saccharide
saccharide
non-polymer
other
```

[chem\_comp]

\* **\_chem\_comp\_angle.atom\_id\_1**

The ID of the first of the three atoms that define the angle. This data item is a pointer to \_chem\_comp\_atom.atom\_id in the CHEM\_COMP\_ATOM category.

\* **\_chem\_comp\_angle.atom\_id\_2**

The ID of the second of the three atoms that define the angle. The second atom is taken to be the apex of the angle. This data item is a pointer to \_chem\_comp\_atom.atom\_id in the CHEM\_COMP\_ATOM category.

\* **\_chem\_comp\_angle.atom\_id\_3**

The ID of the third of the three atoms that define the angle. This data item is a pointer to \_chem\_comp\_atom.atom\_id in the CHEM\_COMP\_ATOM category.

\* **\_chem\_comp\_angle.comp\_id**

This data item is a pointer to \_chem\_comp.id in the CHEM\_COMP category.

**CHEM\_COMP\_ANGLE**

Data items in the CHEM\_COMP\_ANGLE category record details about angles in a chemical component. Angles are designated by three atoms, with the second atom forming the vertex of the angle. Target values may be specified as angles in degrees, as a distance between the first and third atoms, or both.

Category group(s): inclusive\_group

chem\_comp\_group

Category key(s): \_chem\_comp\_angle.comp\_id

\_chem\_comp\_angle.atom\_id\_1

\_chem\_comp\_angle.atom\_id\_2

\_chem\_comp\_angle.atom\_id\_3

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
loop_
_chem_comp_angle.comp_id
_chem_comp_angle.atom_id_1
_chem_comp_angle.atom_id_2
_chem_comp_angle.atom_id_3
_chem_comp_angle.value_angle
_chem_comp_angle.value_dist
phe N CA C xxx.xxx x.xxx
phe CA C O xxx.xxx x.xxx
phe CB CA C xxx.xxx x.xxx
phe CB CA N xxx.xxx x.xxx
phe CA CB CG xxx.xxx x.xxx
phe CB CG CD1 xxx.xxx x.xxx
phe CB CG CD2 xxx.xxx x.xxx
phe CD1 CG CD2 xxx.xxx x.xxx
phe CG CD1 CE1 xxx.xxx x.xxx
phe CD1 CE1 CZ xxx.xxx x.xxx
phe CE1 CZ CE2 xxx.xxx x.xxx
phe CZ CE2 CD2 xxx.xxx x.xxx
phe CG CD2 CE2 xxx.xxx x.xxx
val N CA C xxx.xxx x.xxx
val CA C O xxx.xxx x.xxx
val CB CA C xxx.xxx x.xxx
val CB CA N xxx.xxx x.xxx
val CA CB CG1 xxx.xxx x.xxx
val CA CB CG2 xxx.xxx x.xxx
val CG1 CB CG2 xxx.xxx x.xxx
```

\_chem\_comp\_angle.value\_angle (float, su)

The value that should be taken as the target value for the angle associated with the specified atoms, expressed in degrees.

The permitted range is [0.0, 180.0].

Related item: \_chem\_comp\_angle.value\_angle\_esd (associated esd).

[chem\_comp\_angle]

\_chem\_comp\_angle.value\_angle\_esd (float)

The standard uncertainty (estimated standard deviation) of \_chem\_comp\_angle.value\_angle.

The permitted range is [0.0, 180.0].

Related item: \_chem\_comp\_angle.value\_angle (associated value).

[chem\_comp\_angle]

\_chem\_comp\_angle.value\_dist (float, su)

The value that should be taken as the target value for the angle associated with the specified atoms, expressed as the distance between the atoms specified by \_chem\_comp\_angle.atom\_id\_1 and \_chem\_comp\_angle.atom\_id\_3.

The permitted range is [0.0, ∞).

Related item: \_chem\_comp\_angle.value\_dist\_esd (associated esd).

[chem\_comp\_angle]

\_chem\_comp\_angle.value\_dist\_esd (float)

The standard uncertainty (estimated standard deviation) of \_chem\_comp\_angle.value\_dist.

The permitted range is [0.0, ∞).

Related item: \_chem\_comp\_angle.value\_dist (associated value).

[chem\_comp\_angle]

## CHEM\_COMP\_ATOM

Data items in the CHEM\_COMP\_ATOM category record details about the atoms in a chemical component. Specifying the atomic coordinates for the components in this category is an alternative to specifying the structure of the component *via* bonds, angles, planes *etc.* in the appropriate CHEM\_COMP subcategories.

Category group(s): `inclusive_group`

`chem_comp_group`

Category key(s): `_chem_comp_atom.comp_id`

`_chem_comp_atom.atom_id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _chem_comp_atom.comp_id
  _chem_comp_atom.atom_id
  _chem_comp_atom.type_symbol
  _chem_comp_atom.substruct_code
  _chem_comp_atom.model_Cartn_x
  _chem_comp_atom.model_Cartn_y
  _chem_comp_atom.model_Cartn_z
  phe N N main 1.20134 0.84658 0.00000
  phe CA C main 0.00000 0.00000 0.00000
  phe C C main -1.25029 0.88107 0.00000
  phe O O main -2.18525 0.66029 -0.78409
  phe CB C side 0.00662 -1.03603 1.11081
  phe CG C side 0.03254 -0.49711 2.50951
  phe CD1 C side -1.15813 -0.12084 3.13467
  phe CE1 C side -1.15720 0.38038 4.42732
  phe CZ C side 0.05385 0.51332 5.11032
  phe CE2 C side 1.26137 0.11613 4.50975
  phe CD2 C side 1.23668 -0.38351 3.20288
  val N N main 1.20134 0.84658 0.00000
  val CA C main 0.00000 0.00000 0.00000
  val C C main -1.25029 0.88107 0.00000
  val O O main -2.18525 0.66029 -0.78409
  val CB C side 0.05260 -0.99339 1.17429
  val CG1 C side -0.13288 -0.31545 2.52668
  val CG2 C side -0.94265 -2.12930 0.99811
```

`_chem_comp_atom.alt_atom_id` (line)

An alternative identifier for the atom. This data item would be used in cases where alternative nomenclatures exist for labelling atoms in a group.

[chem\_comp\_atom]

\* `_chem_comp_atom.atom_id` (atcode)

The value of `_chem_comp_atom.atom_id` must uniquely identify each atom in each monomer in the CHEM\_COMP\_ATOM list. The atom identifiers need not be unique over all atoms in the data block; they need only be unique for each atom in a component. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

```
_atom_site.label_atom_id,
_chem_comp_angle.atom_id_1,
_chem_comp_angle.atom_id_2,
_chem_comp_angle.atom_id_3,
_chem_comp_bond.atom_id_1,
_chem_comp_bond.atom_id_2,
_chem_comp_chir.atom_id,
_chem_comp_chir_atom.atom_id,
_chem_comp_plane_atom.atom_id,
_chem_comp_tor.atom_id_1,
_chem_comp_tor.atom_id_2,
_chem_comp_tor.atom_id_3,
_chem_comp_tor.atom_id_4,
_geom_angle.atom_site_label_atom_id_1,
_geom_angle.atom_site_label_atom_id_2,
_geom_angle.atom_site_label_atom_id_3,
_geom_bond.atom_site_label_atom_id_1,
```

```
_geom_bond.atom_site_label_atom_id_2,
_geom_contact.atom_site_label_atom_id_1,
_geom_contact.atom_site_label_atom_id_2,
_geom_hbond.atom_site_label_atom_id_A,
_geom_hbond.atom_site_label_atom_id_D,
_geom_hbond.atom_site_label_atom_id_H,
_geom_torsion.atom_site_label_atom_id_1,
_geom_torsion.atom_site_label_atom_id_2,
_geom_torsion.atom_site_label_atom_id_3,
_geom_torsion.atom_site_label_atom_id_4,
_struct_conn.ptnr1_label_atom_id,
_struct_conn.ptnr2_label_atom_id,
_struct_sheet_hbond.range_1_beg_label_atom_id,
_struct_sheet_hbond.range_1_end_label_atom_id,
_struct_sheet_hbond.range_2_beg_label_atom_id,
_struct_sheet_hbond.range_2_end_label_atom_id,
_struct_site_gen.label_atom_id. [chem_comp_atom]
```

`_chem_comp_atom.charge` (int)

The net integer charge assigned to this atom. This is the formal charge assignment normally found in chemical diagrams.

The permitted range is [-8, 8]. Where no value is given, the assumed value is '0'.

Examples: '1' (for an ammonium nitrogen), '-1' (for a chloride ion).

[chem\_comp\_atom]

\* `_chem_comp_atom.comp_id`

This data item is a pointer to `_chem_comp.id` in the CHEM\_COMP category.

`_chem_comp_atom.model_Cartn_x` (float, su)

The *x* component of the coordinates for this atom in this component specified as orthogonal ångströms. The choice of reference axis frame for the coordinates is arbitrary. The set of coordinates input for the entity here is intended to correspond to the atomic model used to generate restraints for structure refinement, not to atom sites in the ATOM\_SITE list.

Related item: `_chem_comp_atom.model_Cartn_x_esd` (associated esd).

[chem\_comp\_atom]

`_chem_comp_atom.model_Cartn_x_esd` (float)

The standard uncertainty (estimated standard deviation) of `_chem_comp_atom.model_Cartn_x`.

Related item: `_chem_comp_atom.model_Cartn_x` (associated value).

[chem\_comp\_atom]

`_chem_comp_atom.model_Cartn_y` (float, su)

The *y* component of the coordinates for this atom in this component specified as orthogonal ångströms. The choice of reference axis frame for the coordinates is arbitrary. The set of coordinates input for the entity here is intended to correspond to the atomic model used to generate restraints for structure refinement, not to atom sites in the ATOM\_SITE list.

Related item: `_chem_comp_atom.model_Cartn_y_esd` (associated esd).

[chem\_comp\_atom]

`_chem_comp_atom.model_Cartn_y_esd` (float)

The standard uncertainty (estimated standard deviation) of `_chem_comp_atom.model_Cartn_y`.

Related item: `_chem_comp_atom.model_Cartn_y` (associated value).

[chem\_comp\_atom]

**\_chem\_comp\_atom.model\_Cartn\_z** (float, su)

The z component of the coordinates for this atom in this component specified as orthogonal ångströms. The choice of reference axis frame for the coordinates is arbitrary. The set of coordinates input for the entity here is intended to correspond to the atomic model used to generate restraints for structure refinement, not to atom sites in the ATOM\_SITE list.

Related item: \_chem\_comp\_atom.model\_Cartn\_z\_esd (associated esd).

[chem\_comp\_atom]

**\_chem\_comp\_atom.model\_Cartn\_z\_esd** (float)

The standard uncertainty (estimated standard deviation) of \_chem\_comp\_atom.model\_Cartn\_z.

Related item: \_chem\_comp\_atom.model\_Cartn\_z (associated value).

[chem\_comp\_atom]

**\_chem\_comp\_atom.partial\_charge** (float)

The partial charge assigned to this atom.

[chem\_comp\_atom]

**\_chem\_comp\_atom.substruct\_code** (ucode)

This data item assigns the atom to a substructure of the component, if appropriate.

The data value must be one of the following:

main	main chain of an amino acid
side	side chain of an amino acid
base	base of a nucleic acid
phos	phosphate of a nucleic acid
sugar	sugar of a nucleic acid
none	not appropriate for this monomer

[chem\_comp\_atom]

\* **\_chem\_comp\_atom.type\_symbol**

This data item is a pointer to \_atom\_type.symbol in the ATOM\_TYPE category.

\* **\_chem\_comp\_bond.atom\_id\_1**

The ID of the first of the two atoms that define the bond. This data item is a pointer to \_chem\_comp\_atom.atom\_id in the CHEM\_COMP\_ATOM category.

\* **\_chem\_comp\_bond.atom\_id\_2**

The ID of the second of the two atoms that define the bond. This data item is a pointer to \_chem\_comp\_atom.atom\_id in the CHEM\_COMP\_ATOM category.

\* **\_chem\_comp\_bond.comp\_id**

This data item is a pointer to \_chem\_comp.id in the CHEM\_COMP category.

**\_chem\_comp\_bond.value\_dist** (float, su)

The value that should be taken as the target for the chemical bond associated with the specified atoms, expressed as a distance.

The permitted range is [0.0, ∞).

Related item: \_chem\_comp\_bond.value\_dist\_esd (associated esd).

[chem\_comp\_bond]

**\_chem\_comp\_bond.value\_dist\_esd** (float)

The standard uncertainty (estimated standard deviation) of \_chem\_comp\_bond.value\_dist.

The permitted range is [0.0, ∞).

Related item: \_chem\_comp\_bond.value\_dist (associated value).

[chem\_comp\_bond]

**\_chem\_comp\_bond.value\_order** (ucode)

The value that should be taken as the target for the chemical bond associated with the specified atoms, expressed as a bond order.

The data value must be one of the following:

sing	single bond
doub	double bond
trip	triple bond
quad	quadruple bond
arom	aromatic bond
poly	polymeric bond
delo	delocalized double bond
pi	π bond

Where no value is given, the assumed value is 'sing'.

[chem\_comp\_bond]

### CHEM\_COMP\_BOND

Data items in the CHEM\_COMP\_BOND category record details about the bonds between atoms in a chemical component. Target values may be specified as bond orders, as a distance between the two atoms, or both.

Category group(s): inclusive\_group

chem\_comp\_group

Category key(s): \_chem\_comp\_bond.comp\_id

\_chem\_comp\_bond.atom\_id\_1

\_chem\_comp\_bond.atom\_id\_2

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _chem_comp_bond.comp_id
  _chem_comp_bond.atom_id_1
  _chem_comp_bond.atom_id_2
  _chem_comp_bond.value_order
  phe N CA sing
  phe CA C sing
  phe C O doub
  phe CB CA sing
  phe CB CG sing
  phe CG CD1 arom
  phe CD1 CE1 arom
  phe CE1 CZ arom
  phe CZ CE2 arom
  phe CE2 CD2 arom
  phe CD2 CG arom
  val N CA sing
  val CA C sing
  val C O doub
  val CB CA sing
  val CB CG1 sing
  val CB CG2 sing
```

### CHEM\_COMP\_CHIR

Data items in the CHEM\_COMP\_CHIR category provide details about the chiral centres in a chemical component. The atoms bonded to the chiral atom are specified in the CHEM\_COMP\_CHIR\_ATOM category.

Category group(s): inclusive\_group

chem\_comp\_group

Category key(s): \_chem\_comp\_chir.comp\_id

\_chem\_comp\_chir.id

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _chem_comp_chir.comp_id
  _chem_comp_chir.id
  _chem_comp_chir.atom_id
  phe phe1 CA
  val val1 CA
# - - - data truncated for brevity - - -
```

**chem\_comp\_chir.atom\_config** (ucode)

The chiral configuration of the atom that is a chiral centre.

The data value must be one of the following:

R absolute configuration *R*  
S absolute configuration *S*

[chem\_comp\_chir]

\* **chem\_comp\_chir.atom\_id**

The ID of the atom that is a chiral centre. This data item is a pointer to chem\_comp\_atom.atom\_id in the CHEM\_COMP\_ATOM category.

\* **chem\_comp\_chir.comp\_id**

This data item is a pointer to chem\_comp.id in the CHEM\_COMP category.

\* **chem\_comp\_chir.id** (code)

The value of chem\_comp\_chir.id must uniquely identify a record in the CHEM\_COMP\_CHIR list.

The following item(s) have an equivalent role in their respective categories:

chem\_comp\_chir\_atom.chir\_id [chem\_comp\_chir]

**chem\_comp\_chir.number\_atoms\_all** (int)

The total number of atoms bonded to the atom specified by chem\_comp\_chir.atom\_id.

[chem\_comp\_chir]

**chem\_comp\_chir.number\_atoms\_nh** (int)

The number of non-hydrogen atoms bonded to the atom specified by chem\_comp\_chir.atom\_id.

[chem\_comp\_chir]

**chem\_comp\_chir.volume\_flag** (ucode)

A flag to indicate whether a chiral volume should match the standard value in both magnitude and sign, or in magnitude only.

The data value must be one of the following:

sign match magnitude and sign  
nosign match magnitude only

[chem\_comp\_chir]

**chem\_comp\_chir.volume\_three** (float, su)

The chiral volume,  $V_c$ , for chiral centres that involve a chiral atom bonded to three non-hydrogen atoms and one hydrogen atom.

$$V_c = \mathbf{V}_1 \cdot (\mathbf{V}_2 \times \mathbf{V}_3),$$

where  $\mathbf{V}_1$  = the vector distance from the atom specified by chem\_comp\_chir.atom\_id to the first atom in the CHEM\_COMP\_CHIR\_ATOM list,  $\mathbf{V}_2$  = the vector distance from the atom specified by chem\_comp\_chir.atom\_id to the second atom in the CHEM\_COMP\_CHIR\_ATOM list,  $\mathbf{V}_3$  = the vector distance from the atom specified by chem\_comp\_chir.atom\_id to the third atom in the CHEM\_COMP\_CHIR\_ATOM list,  $\cdot$  = the vector dot product and  $\times$  = the vector cross product.

Related item: chem\_comp\_chir.volume\_three\_esd (associated esd).

[chem\_comp\_chir]

**chem\_comp\_chir.volume\_three\_esd** (float)

The standard uncertainty (estimated standard deviation) of chem\_comp\_chir.volume\_three.

Related item: chem\_comp\_chir.volume\_three (associated value).

[chem\_comp\_chir]

### CHEM\_COMP\_CHIR\_ATOM

Data items in the CHEM\_COMP\_CHIR\_ATOM category enumerate the atoms bonded to a chiral atom within a chemical component.

Category group(s): inclusive\_group

chem\_comp\_group

Category key(s): chem\_comp\_chir\_atom.chir\_id

chem\_comp\_chir\_atom.atom\_id

chem\_comp\_chir\_atom.comp\_id

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
loop_
  _chem_comp_chir_atom.comp_id
  _chem_comp_chir_atom.chir_id
  _chem_comp_chir_atom.atom_id
  phe 1 N
  phe 1 C
  phe 1 CB
  val 1 N
  val 1 C
  val 1 CB
  # - - - data truncated for brevity - - -
```

\* **chem\_comp\_chir\_atom.atom\_id**

The ID of an atom bonded to the chiral atom.

This data item is a pointer to chem\_comp\_atom.atom\_id in the CHEM\_COMP\_ATOM category.

\* **chem\_comp\_chir\_atom.chir\_id**

This data item is a pointer to chem\_comp\_chir.id in the CHEM\_COMP\_CHIR category.

\* **chem\_comp\_chir\_atom.comp\_id**

This data item is a pointer to chem\_comp.id in the CHEM\_COMP category.

**chem\_comp\_chir\_atom.dev** (float)

The standard uncertainty (estimated standard deviation) of the position of this atom from the plane defined by all of the atoms in the plane.

[chem\_comp\_chir\_atom]

### CHEM\_COMP\_LINK

Data items in the CHEM\_COMP\_LINK category give details about the links between chemical components.

Category group(s): inclusive\_group

chem\_link\_group

Category key(s): chem\_comp\_link.link\_id

**chem\_comp\_link.details** (text)

A description of special aspects of a link between chemical components in the structure.

[chem\_comp\_link]

\* **chem\_comp\_link.link\_id**

This data item is a pointer to chem\_link.id in the CHEM\_LINK category.

\* **chem\_comp\_link.type\_comp\_1**

The type of the first of the two components joined by the link. This data item is a pointer to chem\_comp.type in the CHEM\_COMP category.

## \*\_chem\_comp\_link.type\_comp\_2

The type of the second of the two components joined by the link. This data item is a pointer to `_chem_comp.type` in the CHEM\_COMP category.

**CHEM.COMP.PLANE**

Data items in the CHEM\_COMP\_PLANE category provide identifiers for the planes in a chemical component. The atoms in the plane are specified in the CHEM\_COMP\_PLANE\_ATOM category.

Category group(s): `inclusive_group`  
`chem_comp_group`  
 Category key(s): `_chem_comp_plane.comp_id`  
`_chem_comp_plane.id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _chem_comp_plane.comp_id
  _chem_comp_plane.id
  phe phe1
```

## \*\_chem\_comp\_plane.comp\_id

This data item is a pointer to `_chem_comp.id` in the CHEM\_COMP category.

## \*\_chem\_comp\_plane.id (code)

The value of `_chem_comp_plane.id` must uniquely identify a record in the CHEM\_COMP\_PLANE list.

The following item(s) have an equivalent role in their respective categories:

`_chem_comp_plane_atom.plane_id` [chem\_comp\_plane]

`_chem_comp_plane.number_atoms_all` (int)

The total number of atoms in the plane.

[chem\_comp\_plane]

`_chem_comp_plane.number_atoms_nh` (int)

The number of non-hydrogen atoms in the plane.

[chem\_comp\_plane]

**CHEM.COMP.PLANE.ATOM**

Data items in the CHEM\_COMP\_PLANE\_ATOM category enumerate the atoms in a plane within a chemical component.

Category group(s): `inclusive_group`  
`chem_comp_group`  
 Category key(s): `_chem_comp_plane_atom.plane_id`  
`_chem_comp_plane_atom.atom_id`  
`_chem_comp_plane_atom.comp_id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _chem_comp_plane_atom.plane_id
  _chem_comp_plane_atom.comp_id
  _chem_comp_plane_atom.atom_id
  phe1 phe CB
  phe1 phe CG
  phe1 phe CD1
  phe1 phe CE1
  phe1 phe CZ
  phe1 phe CE2
  phe1 phe CD2
```

## \*\_chem\_comp\_plane\_atom.atom\_id

The ID of an atom involved in the plane. This data item is a pointer to `_chem_comp_atom.atom_id` in the CHEM\_COMP\_ATOM category.

## \*\_chem\_comp\_plane\_atom.comp\_id

This data item is a pointer to `_chem_comp.id` in the CHEM\_COMP category.

`_chem_comp_plane_atom.dist_esd` (float)

This data item is the standard deviation of the out-of-plane distance for this atom.

## \*\_chem\_comp\_plane\_atom.plane\_id

This data item is a pointer to `_chem_comp_plane.id` in the CHEM\_COMP\_PLANE category.

**CHEM.COMP.TOR**

Data items in the CHEM\_COMP\_TOR category record details about the torsion angles in a chemical component. As torsion angles can have more than one target value, the target values are specified in the CHEM\_COMP\_TOR\_VALUE category.

Category group(s): `inclusive_group`  
`chem_comp_group`  
 Category key(s): `_chem_comp_tor.comp_id`  
`_chem_comp_tor.id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _chem_comp_tor.comp_id
  _chem_comp_tor.id
  _chem_comp_tor.atom_id_1
  _chem_comp_tor.atom_id_2
  _chem_comp_tor.atom_id_3
  _chem_comp_tor.atom_id_4
  phe phe_chi1 N CA CB CG
  phe phe_chi2 CA CB CG CD1
  phe phe_ring1 CB CG CD1 CE1
  phe phe_ring2 CB CG CD2 CE2
  phe phe_ring3 CG CD1 CE1 CZ
  phe phe_ring4 CD1 CE1 CZ CE2
  phe phe_ring5 CE1 CZ CE2 CD2
```

## \*\_chem\_comp\_tor.atom\_id\_1

The ID of the first of the four atoms that define the torsion angle. This data item is a pointer to `_chem_comp_atom.atom_id` in the CHEM\_COMP\_ATOM category.

## \*\_chem\_comp\_tor.atom\_id\_2

The ID of the second of the four atoms that define the torsion angle. This data item is a pointer to `_chem_comp_atom.atom_id` in the CHEM\_COMP\_ATOM category.

## \*\_chem\_comp\_tor.atom\_id\_3

The ID of the third of the four atoms that define the torsion angle. This data item is a pointer to `_chem_comp_atom.atom_id` in the CHEM\_COMP\_ATOM category.

## \*\_chem\_comp\_tor.atom\_id\_4

The ID of the fourth of the four atoms that define the torsion angle. This data item is a pointer to `_chem_comp_atom.atom_id` in the CHEM\_COMP\_ATOM category.

## \*\_chem\_comp\_tor.comp\_id

This data item is a pointer to `_chem_comp.id` in the CHEM\_COMP category.

\* **`_chem_comp_tor.id`** (code)

The value of `_chem_comp_tor.id` must uniquely identify a record in the CHEM\_COMP\_TOR list.

The following item(s) have an equivalent role in their respective categories:

`_chem_comp_tor_value.tor_id` [chem\_comp\_tor]

### CHEM\_COMP\_TOR\_VALUE

Data items in the CHEM\_COMP\_TOR\_VALUE category record details about the target values for the torsion angles enumerated in the CHEM\_COMP\_TOR list. Target values may be specified as angles in degrees, as a distance between the first and fourth atoms, or both.

Category group(s): `inclusive_group`

`chem_comp_group`

Category key(s): `_chem_comp_tor_value.tor_id`

`_chem_comp_tor_value.comp_id`

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
loop_
  _chem_comp_tor_value.tor_id
  _chem_comp_tor_value.comp_id
  _chem_comp_tor_value.angle
  _chem_comp_tor_value.dist
  phe_chi1 phe -60.0 2.88
  phe_chi1 phe 180.0 3.72
  phe_chi1 phe 60.0 2.88
  phe_chi2 phe 90.0 3.34
  phe_chi2 phe -90.0 3.34
  phe_ring1 phe 180.0 3.75
  phe_ring2 phe 180.0 3.75
  phe_ring3 phe 0.0 2.80
  phe_ring4 phe 0.0 2.80
  phe_ring5 phe 0.0 2.80
```

\* **`_chem_comp_tor_value.angle`** (float, su)

A value that should be taken as a potential target value for the torsion angle associated with the specified atoms, expressed in degrees.

The permitted range is [-180.0, 180.0].

Related item: `_chem_comp_tor_value.angle_esd` (associated esd).

[chem\_comp\_tor\_value]

\* **`_chem_comp_tor_value.angle_esd`** (float)

The standard uncertainty (estimated standard deviation) of `_chem_comp_tor_value.angle`.

The permitted range is [-180.0, 180.0].

Related item: `_chem_comp_tor_value.angle` (associated value).

[chem\_comp\_tor\_value]

\* **`_chem_comp_tor_value.comp_id`**

This data item is a pointer to `_chem_comp_atom.comp_id` in the CHEM\_COMP\_ATOM category.

**`_chem_comp_tor_value.dist`** (float, su)

A value that should be taken as a potential target value for the torsion angle associated with the specified atoms, expressed as the distance between the atoms specified by `_chem_comp_tor.atom_id_1` and `_chem_comp_tor.atom_id_4` in the referenced record in the CHEM\_COMP\_TOR list. Note that the torsion angle cannot be fully specified by a distance (for instance, a torsion angle of  $-60^\circ$  will yield the same distance as a  $60^\circ$  angle). However, the distance specification can be useful for refinement in situations in which the angle is already close to the desired value.

The permitted range is [0.0,  $\infty$ ).

Related item: `_chem_comp_tor_value.dist_esd` (associated esd).

[chem\_comp\_tor\_value]

**`_chem_comp_tor_value.dist_esd`** (float)

The standard uncertainty (estimated standard deviation) of `_chem_comp_tor_value.dist`.

The permitted range is [0.0,  $\infty$ ).

Related item: `_chem_comp_tor_value.dist` (associated value).

[chem\_comp\_tor\_value]

\* **`_chem_comp_tor_value.tor_id`**

This data item is a pointer to `_chem_comp_tor.id` in the CHEM\_COMP\_TOR category.

### CHEMLINK

Data items in the CHEM\_LINK category give details about the links between chemical components.

Category group(s): `inclusive_group`

`chem_link_group`

Category key(s): `_chem_link.id`

**`_chem_link.details`** (text)

A description of special aspects of a link between chemical components in the structure.

[chem\_link]

\* **`_chem_link.id`** (code)

The value of `_chem_link.id` must uniquely identify each item in the CHEM\_LINK list.

The following item(s) have an equivalent role in their respective categories:

`_chem_link_angle.link_id`,

`_chem_link_bond.link_id`,

`_chem_link_chir.link_id`,

`_chem_link_plane.link_id`,

`_chem_link_tor.link_id`,

`_chem_comp_link.link_id`,

`_entity_link.link_id`.

Examples: 'peptide', 'oligosaccharide 1,4', 'DNA'.

[chem\_link]

### CHEM\_LINK\_ANGLE

Data items in the CHEM\_LINK\_ANGLE category record details about angles in a link between chemical components.

Category group(s): `inclusive_group`

`chem_link_group`

Category key(s): `_chem_link_angle.link_id`

`_chem_link_angle.atom_id_1`

`_chem_link_angle.atom_id_2`

`_chem_link_angle.atom_id_3`

Example 1 – Engh & Huber parameters [Acta Cryst. (1991), A47, 392–400] as interpreted by J. P. Priestle (1995). Consistent Stereochemical Dictionaries for Refinement and Model Building. CCP4 Daresbury Study Weekend, DL-CONF-95-001, ISSN 1358-6254. Warrington: Daresbury Laboratory.

```
loop_
  _chem_link_angle.link_id
  _chem_link_angle.value_angle
  _chem_link_angle.value_angle_esd
  _chem_link_angle.atom_id_1
  _chem_link_angle.atom_1_comp_id
  _chem_link_angle.atom_id_2
  _chem_link_angle.atom_2_comp_id
  _chem_link_angle.atom_id_3
  _chem_link_angle.atom_3_comp_id
  PEPTIDE 111.2 2.8 N 1 CA 1 C 1
  PEPTIDE 120.8 1.7 CA 1 C 1 O 1
  PEPTIDE 116.2 2.0 CA 1 C 1 N 2
  PEPTIDE 123.0 1.6 O 1 C 1 N 2
  PEPTIDE 121.7 1.8 C 1 N 2 CA 2
```

**\_chem\_link\_angle.atom\_1\_comp\_id** (ucode)

This data item indicates whether atom 1 is found in the first or the second of the two components connected by the link.

The data value must be one of the following:

- 1 the atom is in component 1
- 2 the atom is in component 2

[chem\_link\_angle]

**\_chem\_link\_angle.atom\_2\_comp\_id** (ucode)

This data item indicates whether atom 2 is found in the first or the second of the two components connected by the link.

The data value must be one of the following:

- 1 the atom is in component 1
- 2 the atom is in component 2

[chem\_link\_angle]

**\_chem\_link\_angle.atom\_3\_comp\_id** (ucode)

This data item indicates whether atom 3 is found in the first or the second of the two components connected by the link.

The data value must be one of the following:

- 1 the atom is in component 1
- 2 the atom is in component 2

[chem\_link\_angle]

**\*\_chem\_link\_angle.atom\_id\_1** (code)

The ID of the first of the three atoms that define the angle. An atom with this ID must exist in the component of the type specified by \_chem\_comp\_link.type\_comp\_1 (or \_chem\_comp\_link.type\_comp\_2, where the appropriate data item is indicated by the value of \_chem\_comp\_angle.atom\_1\_comp\_id).

[chem\_link\_angle]

**\*\_chem\_link\_angle.atom\_id\_2** (code)

The ID of the second of the three atoms that define the angle. The second atom is taken to be the apex of the angle. An atom with this ID must exist in the component of the type specified by \_chem\_comp\_link.type\_comp\_1 (or \_chem\_comp\_link.type\_comp\_2, where the appropriate data item is indicated by the value of \_chem\_comp\_angle.atom\_2\_comp\_id).

[chem\_link\_angle]

**\*\_chem\_link\_angle.atom\_id\_3** (code)

The ID of the third of the three atoms that define the angle. An atom with this ID must exist in the component of the type specified by \_chem\_comp\_link.type\_comp\_1 (or \_chem\_comp\_link.type\_comp\_2, where the appropriate data item is indicated by the value of \_chem\_comp\_angle.atom\_3\_comp\_id).

[chem\_link\_angle]

**\*\_chem\_link\_angle.link\_id**

This data item is a pointer to \_chem\_link.id in the CHEM\_LINK category.

**\_chem\_link\_angle.value\_angle** (float, su)

The value that should be taken as the target value for the angle associated with the specified atoms, expressed in degrees.

The permitted range is [0.0, 180.0].

Related item: \_chem\_link\_angle.value\_angle\_esd (associated esd).

[chem\_link\_angle]

**\_chem\_link\_angle.value\_angle\_esd** (float)

The standard uncertainty (estimated standard deviation) of \_chem\_link\_angle.value\_angle.

The permitted range is [0.0, 180.0].

Related item: \_chem\_link\_angle.value\_angle (associated value).

[chem\_link\_angle]

**\_chem\_link\_angle.value\_dist** (float, su)

The value that should be taken as the target value for the angle associated with the specified atoms, expressed as the distance between the atoms specified by \_chem\_comp\_angle.atom\_id\_1 and \_chem\_comp\_angle.atom\_id\_3.

The permitted range is [0.0, ∞).

Related item: \_chem\_link\_angle.value\_dist\_esd (associated esd).

[chem\_link\_angle]

**\_chem\_link\_angle.value\_dist\_esd** (float)

The standard uncertainty (estimated standard deviation) of \_chem\_comp\_angle.value\_dist.

The permitted range is [0.0, ∞).

Related item: \_chem\_link\_angle.value\_dist (associated value).

[chem\_link\_angle]

### CHEM\_LINK\_BOND

Data items in the CHEM\_LINK\_BOND category record details about bonds in a link between components in the chemical structure.

Category group(s): inclusive\_group  
chem\_link\_group

Category key(s): \_chem\_link\_bond.link\_id  
\_chem\_link\_bond.atom\_id\_1  
\_chem\_link\_bond.atom\_id\_2

*Example 1 – Engh & Huber parameters [Acta Cryst. (1991), A47, 392–400] as interpreted by J. P. Priestle (1995). Consistent Stereochemical Dictionaries for Refinement and Model Building. CCP4 Daresbury Study Weekend, DL-CONF-95-001, ISSN 1358-6254. Warrington: Daresbury Laboratory.*

```
loop_
  _chem_link_bond.link_id
  _chem_link_bond.value_dist
  _chem_link_bond.value_dist_esd
  _chem_link_bond.atom_id_1
  _chem_link_bond.atom_1_comp_id
  _chem_link_bond.atom_id_2
  _chem_link_bond.atom_2_comp_id
  PEPTIDE 1.458 0.019 N 1 CA 1
  PEPTIDE 1.525 0.021 CA 1 C 1
  PEPTIDE 1.329 0.014 C 1 N 2
  PEPTIDE 1.231 0.020 C 1 O 1
```

**\_chem\_link\_bond.atom\_1\_comp\_id** (ucode)

This data item indicates whether atom 1 is found in the first or the second of the two components connected by the link.

The data value must be one of the following:

- 1 the atom is in component 1
- 2 the atom is in component 2

[chem\_link\_bond]

**\_chem\_link\_bond.atom\_2\_comp\_id** (ucode)

This data item indicates whether atom 2 is found in the first or the second of the two chemical components connected by the link.

The data value must be one of the following:

- 1 the atom is in component 1
- 2 the atom is in component 2

[chem\_link\_bond]

\* **`_chem_link_bond.atom_id_1`** (code)  
The ID of the first of the two atoms that define the bond. As this data item does not point to a specific atom in a specific chemical component, it is not a child in the linkage sense.

[chem\_link\_bond]

\* **`_chem_link_bond.atom_id_2`** (code)  
The ID of the second of the two atoms that define the bond. As this data item does not point to a specific atom in a specific component, it is not a child in the linkage sense.

[chem\_link\_bond]

\* **`_chem_link_bond.link_id`**  
This data item is a pointer to `_chem_link.id` in the CHEM\_LINK category.

**`_chem_link_bond.value_dist`** (float, su)

The value that should be taken as the target for the chemical bond associated with the specified atoms, expressed as a distance.

The permitted range is [0.0, ∞).

Related item: `_chem_link_bond.value_dist_esd` (associated esd).

[chem\_link\_bond]

**`_chem_link_bond.value_dist_esd`** (float)

The standard uncertainty (estimated standard deviation) of `_chem_link_bond.value_dist`.

The permitted range is [0.0, ∞).

Related item: `_chem_link_bond.value_dist` (associated value).

[chem\_link\_bond]

**`_chem_link_bond.value_order`** (ucode)

The value that should be taken as the target for the chemical bond associated with the specified atoms, expressed as a bond order.

The data value must be one of the following:

sing	single bond
doub	double bond
trip	triple bond
quad	quadruple bond
arom	aromatic bond
poly	polymeric bond
delo	delocalized double bond
pi	π bond

Where no value is given, the assumed value is 'sing'.

[chem\_link\_bond]

### CHEM\_LINK\_CHIR

Data items in the CHEM\_LINK\_CHIR category provide details about the chiral centres in a link between two chemical components. The atoms bonded to the chiral atom are specified in the CHEM\_LINK\_CHIR\_ATOM category.

Category group(s): `inclusive_group`  
`chem_link_group`

Category key(s): `chem_link_chir.link_id`  
`_chem_link_chir.id`

**`_chem_link_chir.atom_comp_id`** (ucode)

This data item indicates whether the chiral atom is found in the first or the second of the two components connected by the link.

The data value must be one of the following:

1	the atom is in component 1
2	the atom is in component 2

[chem\_link\_chir]

**`_chem_link_chir.atom_config`** (ucode)

The chiral configuration of the atom that is a chiral centre.

The data value must be one of the following:

R	absolute configuration <i>R</i>
S	absolute configuration <i>S</i>

[chem\_link\_chir]

\* **`_chem_link_chir.atom_id`** (code)

The ID of the atom that is a chiral centre. As this data item does not point to a specific atom in a specific chemical component, it is not a child in the linkage sense.

\* **`_chem_link_chir.id`** (code)

The value of `_chem_link_chir.id` must uniquely identify a record in the CHEM\_LINK\_CHIR list.

The following item(s) have an equivalent role in their respective categories:

`_chem_link_chir_atom.chir_id`. [chem\_link\_chir]

\* **`_chem_link_chir.link_id`**

This data item is a pointer to `_chem_link.id` in the CHEM\_LINK category.

**`_chem_link_chir.number_atoms_all`** (int)

The total number of atoms bonded to the atom specified by `_chem_link_chir.atom_id`.

[chem\_link\_chir]

**`_chem_link_chir.number_atoms_nh`** (int)

The number of non-hydrogen atoms bonded to the atom specified by `_chem_link_chir.atom_id`.

[chem\_link\_chir]

**`_chem_link_chir.volume_flag`** (ucode)

A flag to indicate whether a chiral volume should match the standard value in both magnitude and sign, or in magnitude only.

The data value must be one of the following:

sign	match magnitude and sign
nosign	match magnitude only

[chem\_link\_chir]

**`_chem_link_chir.volume_three`** (float, su)

The chiral volume,  $V_c$ , for chiral centres that involve a chiral atom bonded to three non-hydrogen atoms and one hydrogen atom.

$$V_c = \mathbf{V}_1 \cdot (\mathbf{V}_2 \times \mathbf{V}_3),$$

where  $\mathbf{V}_1$  = the vector distance from the atom specified by `_chem_link_chir.atom_id` to the first atom in the CHEM\_LINK\_CHIR\_ATOM list,  $\mathbf{V}_2$  = the vector distance from the atom specified by `_chem_link_chir.atom_id` to the second atom in the CHEM\_LINK\_CHIR\_ATOM list,  $\mathbf{V}_3$  = the vector distance from the atom specified by `_chem_link_chir.atom_id` to the third atom in the CHEM\_LINK\_CHIR\_ATOM list,  $\cdot$  = the vector dot product and  $\times$  = the vector cross product.

Related item: `_chem_link_chir.volume_three_esd` (associated esd).

[chem\_link\_chir]

**`_chem_link_chir.volume_three_esd`** (float)

The standard uncertainty (estimated standard deviation) of `_chem_link_chir.volume_three`.

Related item: `_chem_link_chir.volume_three` (associated value).

[chem\_link\_chir]

**CHEM\_LINK\_CHIR\_ATOM**

Data items in the CHEM\_LINK\_CHIR\_ATOM category enumerate the atoms bonded to a chiral atom in a link between two chemical components.

Category group(s): **inclusive\_group**  
                           **chem\_link\_group**  
 Category key(s): **\_chem\_link\_chir\_atom.chir\_id**  
                           **\_chem\_link\_chir\_atom.atom\_id**

**\_chem\_link\_chir\_atom.atom\_comp\_id** (ucode)

This data item indicates whether the atom bonded to a chiral atom is found in the first or the second of the two components connected by the link.

The data value must be one of the following:

- 1 the atom is in component 1
- 2 the atom is in component 2

[chem\_link\_chir\_atom]

\* **\_chem\_link\_chir\_atom.atom\_id** (code)

The ID of an atom bonded to the chiral atom. As this data item does not point to a specific atom in a specific chemical component, it is not a child in the linkage sense.

\* **\_chem\_link\_chir\_atom.chir\_id**

This data item is a pointer to **\_chem\_link\_chir.id** in the CHEM\_LINK\_CHIR category.

**\_chem\_link\_chir\_atom.dev** (float)

The standard uncertainty (estimated standard deviation) of the position of this atom from the plane defined by all of the atoms in the plane.

[chem\_link\_chir\_atom]

**CHEM\_LINK\_PLANE**

Data items in the CHEM\_LINK\_PLANE category provide identifiers for the planes in a link between two chemical components. The atoms in the plane are specified in the CHEM\_LINK\_PLANE\_ATOM category.

Category group(s): **inclusive\_group**  
                           **chem\_link\_group**  
 Category key(s): **\_chem\_link\_plane.link\_id**  
                           **\_chem\_link\_plane.id**

\* **\_chem\_link\_plane.id** (code)

The value of **\_chem\_link\_plane.id** must uniquely identify a record in the CHEM\_LINK\_PLANE list.

The following item(s) have an equivalent role in their respective categories:

**\_chem\_link\_plane\_atom.plane\_id** [chem\_link\_plane]

\* **\_chem\_link\_plane.link\_id**

This data item is a pointer to **\_chem\_link.id** in the CHEM\_LINK category.

**\_chem\_link\_plane.number\_atoms\_all** (int)

The total number of atoms in the plane.

[chem\_link\_plane]

**\_chem\_link\_plane.number\_atoms\_nh** (int)

The number of non-hydrogen atoms in the plane.

[chem\_link\_plane]

**CHEM\_LINK\_PLANE\_ATOM**

Data items in the CHEM\_LINK\_PLANE\_ATOM category enumerate the atoms in a plane in a link between two chemical components.

Category group(s): **inclusive\_group**  
                           **chem\_link\_group**  
 Category key(s): **\_chem\_link\_plane\_atom.plane\_id**  
                           **\_chem\_link\_plane\_atom.atom\_id**

**\_chem\_link\_plane\_atom.atom\_comp\_id** (ucode)

This data item indicates whether the atom in a plane is found in the first or the second of the two components connected by the link.

The data value must be one of the following:

- 1 the atom is in component 1
- 2 the atom is in component 2

[chem\_link\_plane\_atom]

\* **\_chem\_link\_plane\_atom.atom\_id** (code)

The ID of an atom involved in the plane. As this data item does not point to a specific atom in a specific chemical component, it is not a child in the linkage sense.

\* **\_chem\_link\_plane\_atom.plane\_id**

This data item is a pointer to **\_chem\_link\_plane.id** in the CHEM\_LINK\_PLANE category.

**CHEM\_LINK\_TOR**

Data items in the CHEM\_LINK\_TOR category record details about the torsion angles in a link between two chemical components. As torsion angles can have more than one target value, the target values are specified in the CHEM\_LINK\_TOR\_VALUE category.

Category group(s): **inclusive\_group**  
                           **chem\_link\_group**  
 Category key(s): **\_chem\_link\_tor.link\_id**  
                           **\_chem\_link\_tor.id**

**\_chem\_link\_tor.atom\_1\_comp\_id** (ucode)

This data item indicates whether atom 1 is found in the first or the second of the two components connected by the link.

The data value must be one of the following:

- 1 the atom is in component 1
- 2 the atom is in component 2

[chem\_link\_tor]

**\_chem\_link\_tor.atom\_2\_comp\_id** (ucode)

This data item indicates whether atom 2 is found in the first or the second of the two components connected by the link.

The data value must be one of the following:

- 1 the atom is in component 1
- 2 the atom is in component 2

[chem\_link\_tor]

**\_chem\_link\_tor.atom\_3\_comp\_id** (ucode)

This data item indicates whether atom 3 is found in the first or the second of the two components connected by the link.

The data value must be one of the following:

- 1 the atom is in component 1
- 2 the atom is in component 2

[chem\_link\_tor]

**`_chem_link_tor.atom_4_comp_id`** (ucode)

This data item indicates whether atom 4 is found in the first or the second of the two components connected by the link.

The data value must be one of the following:

- 1 the atom is in component 1
- 2 the atom is in component 2

[chem\_link\_tor]

\* **`_chem_link_tor.atom_id_1`** (code)

The ID of the first of the four atoms that define the torsion angle. As this data item does not point to a specific atom in a specific chemical component, it is not a child in the linkage sense.

\* **`_chem_link_tor.atom_id_2`** (code)

The ID of the second of the four atoms that define the torsion angle. As this data item does not point to a specific atom in a specific chemical component, it is not a child in the linkage sense.

\* **`_chem_link_tor.atom_id_3`** (code)

The ID of the third of the four atoms that define the torsion angle. As this data item does not point to a specific atom in a specific chemical component, it is not a child in the linkage sense.

\* **`_chem_link_tor.atom_id_4`** (code)

The ID of the fourth of the four atoms that define the torsion angle. As this data item does not point to a specific atom in a specific chemical component, it is not a child in the linkage sense.

\* **`_chem_link_tor.id`** (code)

The value of `_chem_link_tor.id` must uniquely identify a record in the CHEM\_LINK\_TOR list.

The following item(s) have an equivalent role in their respective categories:

`_chem_link_tor_value.tor_id` [chem\_link\_tor]

\* **`_chem_link_tor.link_id`**

This data item is a pointer to `_chem_link.id` in the CHEM\_LINK category.

### CHEM\_LINK\_TOR\_VALUE

Data items in the CHEM\_LINK\_TOR\_VALUE category record details about the target values for the torsion angles enumerated in the CHEM\_LINK\_TOR list. Target values may be specified as angles in degrees, as a distance between the first and fourth atoms, or both.

Category group(s): `inclusive_group`  
`chem_link_group`

Category key(s): `_chem_link_tor_value.tor_id`

\* **`_chem_link_tor_value.angle`** (float, su)

A value that should be taken as a potential target value for the torsion angle associated with the specified atoms, expressed in degrees.

The permitted range is [-180.0, 180.0].

Related item: `_chem_link_tor_value.angle_esd` (associated esd).

[chem\_link\_tor\_value]

\* **`_chem_link_tor_value.angle_esd`** (float)

The standard uncertainty (estimated standard deviation) of `_chem_link_tor_value.angle`.

The permitted range is [-180.0, 180.0].

Related item: `_chem_link_tor_value.angle` (associated value).

[chem\_link\_tor\_value]

**`_chem_link_tor_value.dist`** (float, su)

A value that should be taken as a potential target value for the torsion angle associated with the specified atoms, expressed as the distance between the atoms specified by `_chem_link_tor.atom_id_1` and `_chem_link_tor.atom_id_4` in the referenced record in the CHEM\_LINK\_TOR list. Note that the torsion angle cannot be fully specified by a distance (for instance, a torsion angle of  $-60^\circ$  will yield the same distance as a  $60^\circ$  angle). However, the distance specification can be useful for refinement in situations in which the angle is already close to the desired value.

The permitted range is [0.0,  $\infty$ ).

Related item: `_chem_link_tor_value.dist_esd` (associated esd).

[chem\_link\_tor\_value]

**`_chem_link_tor_value.dist_esd`** (float)

The standard uncertainty (estimated standard deviation) of `_chem_link_tor_value.dist`.

The permitted range is [0.0,  $\infty$ ).

Related item: `_chem_link_tor_value.dist` (associated value).

[chem\_link\_tor\_value]

\* **`_chem_link_tor_value.tor_id`**

This data item is a pointer to `_chem_link_tor.id` in the CHEM\_LINK\_TOR category.

### CHEMICAL

Data items in the CHEMICAL category would not in general be used in a macromolecular CIF. See instead the ENTITY data items. Data items in the CHEMICAL category record details about the composition and chemical properties of the compounds. The formula data items must agree with those that specify the density, unit-cell and Z values.

Category group(s): `inclusive_group`  
`chemical_group`

Category key(s): `_chemical.entry_id`

Example 1 – based on data set 9597gaus of Alyea, Ferguson & Kannan [Acta Cryst. (1996), C52, 765–767].

```
_chemical.entry_id      '9597gaus'
_chemical.name_systematic
trans-bis(tricyclohexylphosphine)tetracarbonylmolybdenum(0)
```

**`_chemical.absolute_configuration`** (code)

`_chemical.absolute_configuration` (cif\_core.dic 2.3)

Necessary conditions for the assignment of `_chemical.absolute_configuration` are given by H. D. Flack and G. Bernardinelli (1999, 2000).

References: Flack, H. D. & Bernardinelli, G. (1999). *Acta Cryst.* **A55**, 908–915. Flack, H. D. & Bernardinelli, G. (2000). *J. Appl. Cryst.* **33**, 1143–1148.

The data value must be one of the following:

- `rm` Absolute configuration established by the structure determination of a compound containing a chiral reference molecule of known absolute configuration.
- `ad` Absolute configuration established by anomalous-dispersion effects in diffraction measurements on the crystal.
- `rmad` Absolute configuration established by the structure determination of a compound containing a chiral reference molecule of known absolute configuration and confirmed by anomalous-dispersion effects in diffraction measurements on the crystal.
- `syn` Absolute configuration has not been established by anomalous-dispersion effects in diffraction measurements on the crystal. The enantiomer has been assigned by reference to an unchanging chiral centre in the synthetic procedure.

**unk** Absolute configuration is unknown, there being no firm chemical evidence for its assignment to hand and it having not been established by anomalous-dispersion effects in diffraction measurements on the crystal. An arbitrary choice of enantiomer has been made.

Inapplicable.

[chemical]

**\_chemical.compound\_source** (text)

*\_chemical\_compound\_source* (cif\_core.dic 2.0.1)

Description of the source of the compound under study, or of the parent molecule if a simple derivative is studied. This includes the place of discovery for minerals or the actual source of a natural product.

Examples: 'From Norilsk (USSR)',

'Extracted from the bark of Cinchona Naturalis'.

[chemical]

\* **\_chemical.entry\_id**

This data item is a pointer to **\_entry.id** in the ENTRY category.

**\_chemical.melting\_point** (float)

*\_chemical\_melting\_point* (cif\_core.dic 2.0.1)

The temperature in kelvins at which the crystalline solid changes to a liquid.

The permitted range is [0.0, ∞).

[chemical]

**\_chemical.melting\_point\_gt** (float)

*\_chemical\_melting\_point\_gt* (cif\_core.dic 2.3)

A temperature in kelvins above which the melting point (the temperature at which the crystalline solid changes to a liquid) lies. **\_chemical.melting\_point\_gt** and **\_chemical.melting\_point\_lt** allow a range of temperatures to be given. **\_chemical.melting\_point** should always be used in preference to these two items whenever possible.

The permitted range is [0.0, ∞).

Related item: **\_chemical.melting\_point** (alternate).

[chemical]

**\_chemical.melting\_point\_lt** (float)

*\_chemical\_melting\_point\_lt* (cif\_core.dic 2.3)

A temperature in kelvins below which the melting point (the temperature at which the crystalline solid changes to a liquid) lies. **\_chemical.melting\_point\_gt** and **\_chemical.melting\_point\_lt** allow a range of temperatures to be given. **\_chemical.melting\_point** should always be used in preference to these two items whenever possible.

The permitted range is [0.0, ∞).

Related item: **\_chemical.melting\_point** (alternate).

[chemical]

**\_chemical.name\_common** (text)

*\_chemical\_name\_common* (cif\_core.dic 2.0.1)

Trivial name by which the compound is commonly known.

Example: '1-bromoestradiol'.

[chemical]

**\_chemical.name\_mineral** (text)

*\_chemical\_name\_mineral* (cif\_core.dic 2.0.1)

Mineral name accepted by the International Mineralogical Association. Use only for natural minerals. See also **\_chemical.compound\_source**.

Example: 'chalcopyrite'.

[chemical]

**\_chemical.name\_structure\_type** (text)

*\_chemical\_name\_structure\_type* (cif\_core.dic 2.0.1)

Commonly used structure-type name. Usually only applied to minerals or inorganic compounds.

Examples: 'perovskite', 'sphalerite', 'A15'.

[chemical]

**\_chemical.name\_systematic** (text)

*\_chemical\_name\_systematic* (cif\_core.dic 2.0.1)

IUPAC or *Chemical Abstracts* full name of the compound.

Example: '1-bromoestra-1,3,5(10)-triene-3,17\b-diol'. [chemical]

**\_chemical.optical\_rotation** (line)

*\_chemical\_optical\_rotation* (cif\_core.dic 2.3)

The optical rotation in solution of the compound is specified in the following format:

$$[\alpha]_{\text{WAVE}}^{\text{TEMP}} = \text{SORT} \quad (c = \text{CONC, SOLV}),$$

where TEMP is the temperature of the measurement in degrees Celsius, WAVE is an indication of the wavelength of the light used for the measurement, CONC is the concentration of the solution given as the mass of the substance in g per 100 ml of solution, SORT is the signed value (preceded by a + or a - sign) of  $100\alpha/(lc)$ , where  $\alpha$  is the signed optical rotation in degrees measured in a cell of length  $l$  in dm and  $c$  is the value of CONC as defined above, and SOLV is the chemical formula of the solvent.

Example: '\[a]^25~D~ = +108 (c = 3.42, CHCl~3~)'. [chemical]

**\_chemical.properties\_biological** (text)

*\_chemical\_properties\_biological* (cif\_core.dic 2.3)

A free-text description of the biological properties of the material.

Examples:

; diverse biological activities including use as a laxative and strong antibacterial activity against *S. aureus* and weak activity against cyclooxygenase-1 (COX-1)

; antibiotic activity against *Bacillus subtilis* (ATCC 6051) but no significant activity against *Candida albicans* (ATCC 14053), *Aspergillus flavus* (NRRL 6541) and *Fusarium verticillioides* (NRRL 25457)

; weakly potent lipoxygenase nonredox inhibitor

; no influenza A virus sialidase inhibitory and plaque reduction activities

; low toxicity against *Drosophila melanogaster*

[chemical]

**\_chemical.properties\_physical** (text)

*\_chemical\_properties\_physical* (cif\_core.dic 2.3)

A free-text description of the physical properties of the material.

Examples: 'air-sensitive', 'moisture-sensitive', 'hygroscopic', 'deliquescent', 'oxygen-sensitive', 'photo-sensitive', 'pyrophoric', 'semiconductor', 'ferromagnetic at low temperature', 'paramagnetic and thermochromic'.

[chemical]

**\_chemical.temperature\_decomposition** (float, su)

*\_chemical\_temperature\_decomposition* (cif\_core.dic 2.3)

The temperature in kelvins at which the solid decomposes.

The permitted range is [0.0, ∞).

Related item: **\_chemical.temperature\_decomposition\_esd** (associated esd).

Example: '350'.

[chemical]

**\_chemical.temperature\_decomposition\_esd** (float)

The estimated standard deviation of **\_chemical.temperature\_decomposition**.

Related item: **\_chemical.temperature\_decomposition** (associated value).

[chemical]

**`_chemical.temperature_decomposition_gt`** (float)  
`_chemical_temperature_decomposition_gt` (cif.core.dic 2.3)

A temperature in kelvins above which the solid is known to decompose. `_chemical.temperature_decomposition_gt` and `_chemical.temperature_decomposition_lt` allow a range of temperatures to be given. `_chemical.temperature_decomposition` should always be used in preference to these two items whenever possible.

The permitted range is [0.0, ∞).

Related item: `_chemical.temperature_decomposition` (alternate).

Example: '350'. [chemical]

**`_chemical.temperature_decomposition_lt`** (float)  
`_chemical_temperature_decomposition_lt` (cif.core.dic 2.3)

A temperature in kelvins below which the solid is known to decompose. `_chemical.temperature_decomposition_gt` and `_chemical.temperature_decomposition_lt` allow a range of temperatures to be given. `_chemical.temperature_decomposition` should always be used in preference to these two items whenever possible.

The permitted range is [0.0, ∞).

Related item: `_chemical.temperature_decomposition` (alternate).

Example: '350'. [chemical]

**`_chemical.temperature_sublimation`** (float, su)  
`_chemical_temperature_sublimation` (cif.core.dic 2.3)

The temperature in kelvins at which the solid sublimes.

The permitted range is [0.0, ∞).

Related item: `_chemical.temperature_sublimation_esd` (associated esd).

Example: '350'. [chemical]

**`_chemical.temperature_sublimation_esd`** (float)  
 The estimated standard deviation of `_chemical.temperature_sublimation`.

Related item: `_chemical.temperature_sublimation` (associated value).

[chemical]

**`_chemical.temperature_sublimation_gt`** (float)  
`_chemical_temperature_sublimation_gt` (cif.core.dic 2.3)

A temperature in kelvins above which the solid is known to sublime. `_chemical.temperature_sublimation_gt` and `_chemical.temperature_sublimation_lt` allow a range of temperatures to be given. `_chemical.temperature_sublimation` should always be used in preference to these two items whenever possible.

The permitted range is [0.0, ∞).

Related item: `_chemical.temperature_sublimation` (alternate).

Example: '350'. [chemical]

**`_chemical.temperature_sublimation_lt`** (float)  
`_chemical_temperature_sublimation_lt` (cif.core.dic 2.3)

A temperature in kelvins below which the solid is known to sublime. `_chemical.temperature_sublimation_gt` and `_chemical.temperature_sublimation_lt` allow a range of temperatures to be given. `_chemical.temperature_sublimation` should always be used in preference to these two items whenever possible.

The permitted range is [0.0, ∞).

Related item: `_chemical.temperature_sublimation` (alternate).

Example: '350'. [chemical]

### CHEMICAL\_CONN\_ATOM

Data items in the CHEMICAL\_CONN\_ATOM category would not, in general, be used in a macromolecular CIF. See instead the ENTITY data items. Data items in the CHEMICAL\_CONN\_ATOM and CHEMICAL\_CONN\_BOND categories record details about the two-dimensional (2D) chemical structure of the molecular species. They allow a 2D chemical diagram to be reconstructed for use in a publication or in a database search for structural and sub-structural relationships. The CHEMICAL\_CONN\_ATOM data items provide information about the chemical properties of the atoms in the structure. In cases where crystallographic and molecular symmetry elements coincide, they must also contain symmetry-generated atoms, so that the CHEMICAL\_CONN\_ATOM and CHEMICAL\_CONN\_BOND data items will always describe a complete chemical entity.

Category group(s): `inclusive_group`  
`chemical_group`

Category key(s): `_chemical_conn_atom.number`

Example 1 – based on data set DPTD of Yamin, Suwandi, Fun, Sivakumar & bin Shawkataly [Acta Cryst. (1996), C52, 951–953].

```
loop_
  _chemical_conn_atom.number
  _chemical_conn_atom.type_symbol
  _chemical_conn_atom.display_x
  _chemical_conn_atom.display_y
  _chemical_conn_atom.NCA
  _chemical_conn_atom.NH
  1 S .39 .81 1 0
  2 S .39 .96 2 0
  3 N .14 .88 3 0
  4 C .33 .88 3 0
  5 C .11 .96 2 2
  6 C .03 .96 2 2
  7 C .03 .80 2 2
  8 C .11 .80 2 2
  9 S .54 .81 1 0
  10 S .54 .96 2 0
  11 N .80 .88 3 0
  12 C .60 .88 3 0
  13 C .84 .96 2 2
  14 C .91 .96 2 2
  15 C .91 .80 2 2
  16 C .84 .80 2 2
```

**`_chemical_conn_atom.charge`** (int)

`_chemical_conn_atom_charge` (cif.core.dic 2.0.1)

The net integer charge assigned to this atom. This is the formal charge assignment normally found in chemical diagrams.

The permitted range is [−8, 8]. Where no value is given, the assumed value is '0'.

Examples: '1' (for an ammonium nitrogen), '-1' (for a chloride ion).

[chemical\_conn\_atom]

**`_chemical_conn_atom.display_x`** (float)

`_chemical_conn_atom_display_x` (cif.core.dic 2.0.1)

The 2D Cartesian *x* coordinate of the position of this atom in a recognizable chemical diagram. The coordinate origin is at the lower left corner, the *x* axis is horizontal and the *y* axis is vertical. The coordinates must lie in the range 0.0 to 1.0. These coordinates can be obtained from projections of a suitable uncluttered view of the molecular structure.

The permitted range is [0.0, 1.0].

[chemical\_conn\_atom]

**`_chemical_conn_atom.display_y`** (float)

`_chemical_conn_atom_display_y` (cif.core.dic 2.0.1)

The 2D Cartesian *y* coordinate of the position of this atom in a recognizable chemical diagram. The coordinate origin is at the lower left corner, the *x* axis is horizontal and the *y* axis is vertical. The coordinates must lie in the range 0.0 to 1.0. These coordinates can be obtained from projections of a suitable uncluttered view of the molecular structure.

The permitted range is [0.0, 1.0].

[chemical\_conn\_atom]

**\_chemical\_conn\_atom.NCA** (int)  
 \_chemical\_conn\_atom\_NCA (cif\_core.dic 2.0.1)

The number of connected atoms excluding terminal hydrogen atoms.

The permitted range is [0, ∞). [chemical\_conn\_atom]

**\_chemical\_conn\_atom.NH** (int)  
 \_chemical\_conn\_atom\_NH (cif\_core.dic 2.0.1)

The total number of hydrogen atoms attached to this atom, regardless of whether they are included in the refinement or the ATOM\_SITE list. This number is the same as `_atom_site.attached_hydrogens` only if none of the hydrogen atoms appear in the ATOM\_SITE list.

The permitted range is [0, ∞). [chemical\_conn\_atom]

\* **\_chemical\_conn\_atom.number** (int)  
 \_chemical\_conn\_atom\_number (cif\_core.dic 2.0.1)

The chemical sequence number to be associated with this atom. Within an ATOM\_SITE list, this number must match one of the `_atom_site.chemical_conn_number` values.

The following item(s) have an equivalent role in their respective categories:

`_atom_site.chemical_conn_number`,

`_chemical_conn_bond.atom_1`,

`_chemical_conn_bond.atom_2`.

The permitted range is [1, ∞). [chemical\_conn\_atom]

\* **\_chemical\_conn\_atom.type\_symbol**  
 \_chemical\_conn\_atom\_type\_symbol (cif\_core.dic 2.0.1)

This data item is a pointer to `_atom_type.symbol` in the ATOM\_TYPE category.

\* **\_chemical\_conn\_bond.atom\_1**  
 \_chemical\_conn\_bond\_atom\_1 (cif\_core.dic 2.0.1)

This data item is a pointer to `_chemical_conn_atom.number` in the CHEMICAL\_CONN\_ATOM category.

\* **\_chemical\_conn\_bond.atom\_2**  
 \_chemical\_conn\_bond\_atom\_2 (cif\_core.dic 2.0.1)

This data item is a pointer to `_chemical_conn_atom.number` in the CHEMICAL\_CONN\_ATOM category.

**\_chemical\_conn\_bond.type** (ucode)  
 \_chemical\_conn\_bond\_type (cif\_core.dic 2.0.1)

The chemical bond type associated with the connection between the two sites `_chemical_conn_bond.atom_1` and `_chemical_conn_bond.atom_2`.

The data value must be one of the following:

sing	single bond
doub	double bond
trip	triple bond
quad	quadruple bond
arom	aromatic bond
poly	polymeric bond
delo	delocalized double bond
pi	π bond

Where no value is given, the assumed value is 'sing'. [chemical\_conn\_bond]

### CHEMICAL\_CONN\_BOND

Data items in the CHEMICAL\_CONN\_BOND category would not, in general, be used in a macromolecular CIF. See instead the ENTITY data items. Data items in the CHEMICAL\_CONN\_ATOM and CHEMICAL\_CONN\_BOND categories record details about the two-dimensional (2D) chemical structure of the molecular species. They allow a 2D chemical diagram to be reconstructed for use in a publication or in a database search for structural and substructural relationships. The CHEMICAL\_CONN\_BOND data items specify the connections between the atoms in the CHEMICAL\_CONN\_ATOM list and the nature of the chemical bond between these atoms.

Category group(s): `inclusive_group`  
`chemical_group`

Category key(s): `_chemical_conn_bond.atom_1`  
`_chemical_conn_bond.atom_2`

Example 1 – based on data set DPTD of Yamin, Suwandi, Fun, Sivakumar & bin Shawkataly [Acta Cryst. (1996), C52, 951–953].

```
loop_
  _chemical_conn_bond.atom_1
  _chemical_conn_bond.atom_2
  _chemical_conn_bond.type
  4      1      doub      4      3      sing
  4      2      sing       5      3      sing
  6      5      sing       7      6      sing
  8      7      sing       8      3      sing
  10     2      sing      12     9      doub
  12     11     sing      12     10     sing
  13     11     sing      14     13     sing
  15     14     sing      16     15     sing
  16     11     sing      17     5      sing
  18     5      sing      19     6      sing
  20     6      sing      21     7      sing
  22     7      sing      23     8      sing
  24     8      sing      25     13     sing
  26     13     sing      27     14     sing
  28     14     sing      29     15     sing
  30     15     sing      31     16     sing
  32     16     sing
```

### CHEMICAL\_FORMULA

Data items in the CHEMICAL\_FORMULA category would not, in general, be used in a macromolecular CIF. See instead the ENTITY data items. Data items in the CHEMICAL\_FORMULA category specify the composition and chemical properties of the compound. The formula data items must agree with those that specify the density, unit-cell and Z values. The following rules apply to the construction of the data items `_chemical_formula.analytical`, `_chemical_formula.structural` and `_chemical_formula.sum`. For the data item `_chemical_formula.moiety`, the formula construction is broken up into residues or moieties, *i.e.* groups of atoms that form a molecular unit or molecular ion. The rules given below apply within each moiety but different requirements apply to the way that moieties are connected (see `_chemical_formula.moiety`). (1) Only recognized element symbols may be used. (2) Each element symbol is followed by a 'count' number. A count of '1' may be omitted. (3) A space or parenthesis must separate each cluster of (element symbol + count). (4) Where a group of elements is enclosed in parentheses, the multiplier for the group must follow the closing parenthesis. That is, all element and group multipliers are assumed to be printed as subscripted numbers. (An exception to this rule exists for `_chemical_formula.moiety` formulae where pre- and post-multipliers are permitted for molecular units.) (5) Unless the elements are ordered in a manner that corresponds to their chemical structure, as in `_chemical_formula.structural`, the order of the elements within any group or moiety should be: C, then H, then the other elements in alphabetical order of their symbol. This is the 'Hill' system used by *Chemical Abstracts*. This ordering is used in `_chemical_formula.moiety` and `_chemical_formula.sum`.

Category group(s): `inclusive_group`  
`chemical_group`

Category key(s): `_chemical_formula.entry_id`

Example 2 – based on data set TOZ of Willis, Beckwith & Tozer [(1991). Acta Cryst. C47, 2276–2277].

```
_chemical_formula.entry_id      'TOZ'
_chemical_formula.moiety        'C18 H25 N O3'
_chemical_formula.sum           'C18 H25 N O3'
_chemical_formula.weight        303.40
```

**\_chemical\_formula.analytical** (text)  
 \_chemical\_formula\_analytical (cif\_core.dic 2.0.1)

Formula determined by standard chemical analysis including trace elements. See the CHEMICAL\_FORMULA category description for rules for writing chemical formulae. Parentheses are used only for standard uncertainties (estimated standard deviations).

Example: 'Fe2.45 (2) Ni1.60 (3) S4'. [chemical\_formula]

\* **\_chemical\_formula.entry\_id**  
 This data item is a pointer to **\_entry.id** in the ENTRY category.

**\_chemical\_formula.iupac** (text)  
 \_chemical\_formula\_iupac (cif\_core.dic 2.0.1)

Formula expressed in conformance with IUPAC rules for inorganic and metal-organic compounds where these conflict with the rules for any other CHEMICAL\_FORMULA entries. Typically used for formatting a formula in accordance with journal rules. This should appear in the data block in addition to the most appropriate of the other CHEMICAL\_FORMULA data names.

Reference: IUPAC (1990). *Nomenclature of Inorganic Chemistry*. Oxford: Blackwell Scientific Publications.

Example: '[Co Re (C12 H22 P)2 (C O)6].0.5C H3 O H'. [chemical\_formula]

**\_chemical\_formula.moiety** (text)  
 \_chemical\_formula\_moiety (cif\_core.dic 2.0.1)

Formula with each discrete bonded residue or ion shown as a separate moiety. See the CHEMICAL\_FORMULA category description for rules for writing chemical formulae. In addition to the general formulae requirements, the following rules apply: (1) Moieties are separated by commas ','. (2) The order of elements within a moiety follows general rule (5) in the CHEMICAL\_FORMULA category description. (3) Parentheses are not used within moieties but may surround a moiety. Parentheses may not be nested. (4) Charges should be placed at the end of the moiety. The charge '+' or '-' may be preceded by a numerical multiplier and should be separated from the last (element symbol + count) by a space. Pre- or post-multipliers may be used for individual moieties.

Examples: 'C7 H4 Cl Hg N O3 S', 'C12 H17 N4 O S 1+', 'C6 H2 N3 O7 1-', 'C12 H16 N2 O6, 5(H2 O1)', '(Cd 2+)3, (C6 N6 Cr 3-)2, 2(H2 O)'. [chemical\_formula]

**\_chemical\_formula.structural** (text)  
 \_chemical\_formula\_structural (cif\_core.dic 2.0.1)

See the CHEMICAL\_FORMULA category description for the rules for writing chemical formulae for inorganics, organometallics, metal complexes *etc.*, in which bonded groups are preserved as discrete entities within parentheses, with post-multipliers as required. The order of the elements should give as much information as possible about the chemical structure. Parentheses may be used and nested as required. This formula should correspond to the structure as actually reported, *i.e.* trace elements not included in atom-type and atom-site data should not be included in this formula (see also **\_chemical\_formula.analytical**).

Examples: 'Ca ((Cl O3)2 O)2 (H2 O)6', '(Pt (N H3)2 (C5 H7 N3 O)2) (Cl O4)2'. [chemical\_formula]

**\_chemical\_formula.sum** (text)  
 \_chemical\_formula\_sum (cif\_core.dic 2.0.1)

See the CHEMICAL\_FORMULA category description for the rules for writing chemical formulae in which all discrete bonded residues and ions are summed over the constituent elements, following the ordering given in general rule (5) in the CHEMICAL\_FORMULA category description. Parentheses are not normally used.

Example: 'C18 H19 N7 O8 S'. [chemical\_formula]

**\_chemical\_formula.weight** (float)  
 \_chemical\_formula\_weight (cif\_core.dic 2.0.1)

Formula mass in daltons. This mass should correspond to the formulae given under **\_chemical\_formula.structural**, **\_chemical\_formula.moiety** or **\_chemical\_formula.sum** and, together with the Z value and cell parameters, should yield the density given as **\_exptl\_crystal.density\_diffrn**.

The permitted range is [1.0, ∞). [chemical\_formula]

**\_chemical\_formula.weight\_meas** (float)  
 \_chemical\_formula\_weight\_meas (cif\_core.dic 2.0.1)

Formula mass in daltons measured by a non-diffraction experiment.

The permitted range is [1.0, ∞). [chemical\_formula]

### CITATION

Data items in the CITATION category record details about the literature cited as being relevant to the contents of the data block.

Category group(s): **inclusive\_group**  
**citation\_group**

Category key(s): **\_citation.id**

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
loop_
_citation.id
_citation.coordinate_linkage
_citation.title
_citation.country
_citation.journal_abbrev
_citation.journal_volume
_citation.journal_issue
_citation.page_first
_citation.page_last
_citation.year
_citation.journal_id_ASTM
_citation.journal_id_ISSN
_citation.journal_id_CSD
_citation.book_title
_citation.book_publisher
_citation.book_id_ISBN
_citation.details
primary yes
; Crystallographic analysis of a complex between human
immunodeficiency virus type 1 protease and acetyl-pepstatin
at 2.0-Angstroms resolution.
;
US 'J. Biol. Chem.' 265 . 14209 14219 1990
HBCHA3 0021-9258 071 . . .
; The publication that directly relates to this coordinate
set.
;
2 no
; Three-dimensional structure of aspartyl-protease from human
immunodeficiency virus HIV-1.
;
UK 'Nature' 337 . 615 619 1989
NATUAS 0028-0836 006 . . .
; Determination of the structure of the unliganded enzyme.
;
3 no
; Crystallization of the aspartylprotease from human
immunodeficiency virus, HIV-1.
;
US 'J. Biol. Chem.' 264 . 1919 1921 1989
HBCHA3 0021-9258 071 . . .
; Crystallization of the unliganded enzyme.
;
4 no
; Human immunodeficiency virus protease. Bacterial expression
and characterization of the purified aspartic protease.
;
US 'J. Biol. Chem.' 264 . 2307 2312 1989
HBCHA3 0021-9258 071 . . .
; Expression and purification of the enzyme.
;
```

<p><b><code>_citation.abstract</code></b> (text)  <code>_citation_abstract</code> (cif_core.dic 2.0.1)            Abstract for the citation. This is used most when the citation is extracted from a bibliographic database that contains full text or abstract information.            [citation]</p>	<p><b><code>_citation.database_id_CSD</code></b> (code)  <code>_citation_database_id_CSD</code> (cif_core.dic 2.3)            Identifier ('refcode') of the database record in the Cambridge Structural Database that contains details of the cited structure.            Example: 'LEKKUH'.            [citation]</p>
<p><b><code>_citation.abstract_id_CAS</code></b> (text)  <code>_citation_abstract_id_CAS</code> (cif_core.dic 2.0.1)            The <i>Chemical Abstracts</i> Service (CAS) abstract identifier; relevant for journal articles.            [citation]</p>	<p><b><code>_citation.database_id_Medline</code></b> (int)  <code>_citation_database_id_Medline</code> (cif_core.dic 2.0.1)            Accession number used by Medline to categorize a specific bibliographic entry.            The permitted range is [1, ∞).            Example: '89064067'.            [citation]</p>
<p><b><code>_citation.book_id_ISBN</code></b> (line)  <code>_citation_book_id_ISBN</code> (cif_core.dic 2.0.1)            The International Standard Book Number (ISBN) code assigned to the book cited; relevant for books or book chapters.            [citation]</p>	<p><b><code>_citation.details</code></b> (text)  <code>_citation_special_details</code> (cif_core.dic 2.0.1)            A description of special aspects of the relationship of the contents of the data block to the literature item cited.            Examples:            ; citation relates to this precise coordinate set            ;            ; citation relates to earlier low-resolution structure            ;            ; citation relates to further refinement of structure reported in citation 2            ;            [citation]</p>
<p><b><code>_citation.book_publisher</code></b> (text)  <code>_citation_book_publisher</code> (cif_core.dic 2.0.1)            The name of the publisher of the citation; relevant for books or book chapters.            Example: 'John Wiley and Sons'.            [citation]</p>	<p><b><code>*_citation.id</code></b> (code)  <code>_citation_id</code> (cif_core.dic 2.0.1)            The value of <code>_citation.id</code> must uniquely identify a record in the CITATION list. The <code>_citation.id</code> 'primary' should be used to indicate the citation that the author(s) consider to be the most pertinent to the contents of the data block. Note that this item need not be a number; it can be any unique identifier.  <i>The following item(s) have an equivalent role in their respective categories:</i>  <code>_citation_author.citation_id</code>,  <code>_citation_editor.citation_id</code>,  <code>_software.citation_id</code>.            Examples: 'primary', '1', '2'.            [citation]</p>
<p><b><code>_citation.book_publisher_city</code></b> (text)  <code>_citation_book_publisher_city</code> (cif_core.dic 2.0.1)            The location of the publisher of the citation; relevant for books or book chapters.            Example: 'London'.            [citation]</p>	<p><b><code>_citation.journal_abbrev</code></b> (line)  <code>_citation_journal_abbrev</code> (cif_core.dic 2.0.1)            Abbreviated name of the cited journal as given in the <i>Chemical Abstracts Service Source Index</i>.            Example: 'J. Mol. Biol.'.            [citation]</p>
<p><b><code>_citation.book_title</code></b> (text)  <code>_citation_book_title</code> (cif_core.dic 2.0.1)            The title of the book in which the citation appeared; relevant for books or book chapters.            [citation]</p>	<p><b><code>_citation.journal_full</code></b> (text)  <code>_citation_journal_full</code> (cif_core.dic 2.0.1)            Full name of the cited journal; relevant for journal articles.            Example: 'Journal of Molecular Biology'.            [citation]</p>
<p><b><code>_citation.coordinate_linkage</code></b> (ucode)  <code>_citation_coordinate_linkage</code> (cif_core.dic 2.0.1)  <code>_citation.coordinate_linkage</code> states whether this citation is concerned with precisely the set of coordinates given in the data block. If, for instance, the publication described the same structure, but the coordinates had undergone further refinement prior to the creation of the data block, the value of this data item would be 'no'.            The data value must be one of the following:            no citation unrelated to current coordinates            n abbreviation for 'no'            yes citation related to current coordinates            y abbreviation for 'yes'            [citation]</p>	<p><b><code>_citation.journal_id_ASTM</code></b> (line)  <code>_citation_journal_id_ASTM</code> (cif_core.dic 2.0.1)            The American Society for Testing and Materials (ASTM) code assigned to the journal cited (also referred to as the CODEN designator of the <i>Chemical Abstracts</i> Service); relevant for journal articles.            [citation]</p>
<p><b><code>_citation.country</code></b> (line)  <code>_citation_country</code> (cif_core.dic 2.0.1)            The country of publication; relevant for books and book chapters.            [citation]</p>	<p>[citation]</p>

**\_citation.journal\_id\_CSD** (line)

**\_citation\_journal\_id\_CSD** (cif\_core.dic 2.0.1)

The Cambridge Structural Database (CSD) code assigned to the journal cited; relevant for journal articles. This is also the system used at the Protein Data Bank (PDB).

Example: '0070'. [citation]

**\_citation.journal\_id\_ISSN** (line)

**\_citation\_journal\_id\_ISSN** (cif\_core.dic 2.0.1)

The International Standard Serial Number (ISSN) code assigned to the journal cited; relevant for journal articles.

[citation]

**\_citation.journal\_issue** (line)

**\_citation\_journal\_issue** (cif\_core.dic 2.0.1)

Issue number of the journal cited; relevant for journal articles.

Example: '2'. [citation]

**\_citation.journal\_volume** (line)

**\_citation\_journal\_volume** (cif\_core.dic 2.0.1)

Volume number of the journal cited; relevant for journal articles.

Example: '174'. [citation]

**\_citation.language** (line)

**\_citation\_language** (cif\_core.dic 2.0.1)

Language in which the cited article is written.

Example: 'German'. [citation]

**\_citation.page\_first** (line)

**\_citation\_page\_first** (cif\_core.dic 2.0.1)

The first page of the citation; relevant for journal articles, books and book chapters.

[citation]

**\_citation.page\_last** (line)

**\_citation\_page\_last** (cif\_core.dic 2.0.1)

The last page of the citation; relevant for journal articles, books and book chapters.

[citation]

**\_citation.title** (text)

**\_citation\_title** (cif\_core.dic 2.0.1)

The title of the citation; relevant for journal articles, books and book chapters.

Example:

```
; Structure of diferric duck ovotransferrin at 2.35 \xA
resolution.
```

; [citation]

**\_citation.year** (int)

**\_citation\_year** (cif\_core.dic 2.0.1)

The year of the citation; relevant for journal articles, books and book chapters.

Example: '1984'. [citation]

## CITATION\_AUTHOR

Data items in the CITATION\_AUTHOR category record details about the authors associated with the citations in the CITATION list.

Category group(s): **inclusive\_group**  
**citation\_group**

Category key(s): **\_citation\_author.citation\_id**  
**\_citation\_author.name**

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
loop_
_citation_author.citation_id
_citation_author.ordinal
_citation_author.name
primary 1 'Fitzgerald, P.M.D.'
primary 2 'McKeever, B.M.'
primary 3 'Van Middlesworth, J.F.'
primary 4 'Springer, J.P.'
primary 5 'Heimbach, J.C.'
primary 6 'Leu, C.-T.'
primary 7 'Herber, W.K.'
primary 8 'Dixon, R.A.F.'
primary 9 'Darke, P.L.'
2 1 'Navia, M.A.'
2 2 'Fitzgerald, P.M.D.'
2 3 'McKeever, B.M.'
2 4 'Leu, C.-T.'
2 5 'Heimbach, J.C.'
2 6 'Herber, W.K.'
2 7 'Sigal, I.S.'
2 8 'Darke, P.L.'
2 9 'Springer, J.P.'
3 1 'McKeever, B.M.'
3 2 'Navia, M.A.'
3 3 'Fitzgerald, P.M.D.'
3 4 'Springer, J.P.'
3 5 'Leu, C.-T.'
3 6 'Heimbach, J.C.'
3 7 'Herber, W.K.'
3 8 'Sigal, I.S.'
3 9 'Darke, P.L.'
4 1 'Darke, P.L.'
4 2 'Leu, C.-T.'
4 3 'Davis, L.J.'
4 4 'Heimbach, J.C.'
4 5 'Diehl, R.E.'
4 6 'Hill, W.S.'
4 7 'Dixon, R.A.F.'
4 8 'Sigal, I.S.'
```

\* **\_citation\_author.citation\_id**  
**\_citation\_author\_citation\_id** (cif\_core.dic 2.0.1)

This data item is a pointer to **\_citation.id** in the CITATION category.

\* **\_citation\_author.name** (line)  
**\_citation\_author\_name** (cif\_core.dic 2.0.1)

Name of an author of the citation; relevant for journal articles, books and book chapters. The family name(s), followed by a comma and including any dynastic components, precedes the first name(s) or initial(s).

Examples: 'Bleary, Percival R.', 'O'Neil, F.K.', 'Van den Bossche, G.', 'Yang, D.-L.', 'Simonov, Yu.A.'. [citation\_author]

**\_citation\_author.ordinal** (int)  
**\_citation\_author ordinal** (cif\_core.dic 2.0.1)

This data item defines the order of the author's name in the list of authors of a citation.

[citation\_author]

**CITATION\_EDITOR**

Data items in the CITATION\_EDITOR category record details about the editors associated with the books or book chapters cited in the CITATION list.

Category group(s): **inclusive\_group**

**citation\_group**

Category key(s): **\_citation\_editor.citation\_id**

**\_citation\_editor.name**

*Example 1 – hypothetical example.*

loop\_

**\_citation\_editor.citation\_id**

**\_citation\_editor.name**

```
5      'McKeever, B.M.'
5      'Navia, M.A.'
5      'Fitzgerald, P.M.D.'
5      'Springer, J.P.'
```

\* **\_citation\_editor.citation\_id**

**\_citation\_editor\_citation\_id** (cif-core.dic 2.0.1)

This data item is a pointer to **\_citation.id** in the CITATION category.

**\_citation\_editor.name**

(line)

**\_citation\_editor\_name** (cif-core.dic 2.0.1)

Names of an editor of the citation; relevant for books and book chapters. The family name(s), followed by a comma and including any dynastic components, precedes the first name(s) or initial(s).

Examples: 'Bleary, Percival R.', 'O'Neil, F.K.', 'Van den Bossche, G.', 'Yang, D.-L.', 'Simonov, Yu.A.' [citation\_editor]

**\_citation\_editor.ordinal**

(int)

**\_citation\_editor ordinal** (cif-core.dic 2.0.1)

This data item defines the order of the editor's name in the list of editors of a citation.

[citation\_editor]

**COMPUTING**

Data items in the COMPUTING category record details about the computer programs used in the crystal structure analysis. Data items in this category would not, in general, be used in a macromolecular CIF. The category SOFTWARE, which allows a more detailed description of computer programs and their attributes to be given, would be used instead.

Category group(s): **inclusive\_group**

**computing\_group**

Category key(s): **\_computing.entry\_id**

*Example 1 – Rodríguez-Romera, Ruiz-Pérez & Solans [Acta Cryst. (1996), C52, 1415–1417].*

```
_computing.data_collection      'CAD-4 (Enraf-Nonius, 1989)'
_computing.cell_refinement      'CAD-4 (Enraf-Nonius, 1989)'
_computing.data_reduction       'CFEO (Solans, 1978)'
_computing.structure_solution    'SHELXS86 (Sheldrick, 1990)'
_computing.structure_refinement 'SHELXL93 (Sheldrick, 1993)'
_computing.molecular_graphics    'ORTEPII (Johnson, 1976)'
_computing.publication_material 'PARST (Nardelli, 1983)'
```

**\_computing.cell\_refinement**

(text)

**\_computing\_cell\_refinement** (cif-core.dic 2.0.1)

Software used for cell refinement. Give the program or package name and a brief reference.

Example: 'CAD4 (Enraf-Nonius, 1989)'.

[computing]

**\_computing.data\_collection**

(text)

**\_computing\_data\_collection** (cif-core.dic 2.0.1)

Software used for data collection. Give the program or package name and a brief reference.

Example: 'CAD4 (Enraf-Nonius, 1989)'.

[computing]

**\_computing.data\_reduction**

(text)

**\_computing\_data\_reduction** (cif-core.dic 2.0.1)

Software used for data reduction. Give the program or package name and a brief reference.

Example: 'DIFDAT, SORTRF, ADDRIF (Hall & Stewart, 1990)'.

[computing]

\* **\_computing.entry\_id**

This data item is a pointer to **\_entry.id** in the ENTRY category.

**\_computing.molecular\_graphics**

(text)

**\_computing\_molecular\_graphics** (cif-core.dic 2.0.1)

Software used for molecular graphics. Give the program or package name and a brief reference.

Example: 'FRODO (Jones, 1986), ORTEP (Johnson, 1965)'.

[computing]

**\_computing.publication\_material**

(text)

**\_computing\_publication\_material** (cif-core.dic 2.0.1)

Software used for generating material for publication. Give the program or package name and a brief reference.

[computing]

**\_computing.structure\_refinement**

(text)

**\_computing\_structure\_refinement** (cif-core.dic 2.0.1)

Software used for refinement of the structure. Give the program or package name and a brief reference.

Examples: 'SHELX85 (Sheldrick, 1985)', 'X-PLOR (Brunger, 1992)'.

[computing]

**\_computing.structure\_solution**

(text)

**\_computing\_structure\_solution** (cif-core.dic 2.0.1)

Software used for solution of the structure. Give the program or package name and a brief reference.

Example: 'SHELX85 (Sheldrick, 1985)'.

[computing]

**DATABASE**

Data items in the DATABASE category have been superseded by data items in the DATABASE\_2 category. They are included here only for compliance with older CIFs.

Category group(s): **inclusive\_group**

**compliance\_group**

Category key(s): **\_database.entry\_id**

**\_database.code\_CAS**

(line)

**\_database\_code\_CAS** (cif-core.dic 2.3)

The code assigned by *Chemical Abstracts*.

[database]

**\_database.code\_CSD**

(line)

**\_database\_code\_CSD** (cif-core.dic 2.3)

The code assigned by the Cambridge Structural Database.

[database]

**\_database.code\_ICSD**

(line)

**\_database\_code\_ICSD** (cif-core.dic 2.3)

The code assigned by the Inorganic Crystal Structure Database.

[database]

**database.code\_MDF** (line)

database\_code\_MDF (cif\_core.dic 2.3)

The code assigned by the Metals Data File.

[database]

**database.code\_NBS** (line)

database\_code\_NBS (cif\_core.dic 2.3)

The code assigned by the NBS (NIST) Crystal Data Database.

[database]

**database.code\_PDB** (line)

database\_code\_PDB (cif\_core.dic 2.3)

The code assigned by the Protein Data Bank.

[database]

**database.code\_PDF** (line)

database\_code\_PDF (cif\_core.dic 2.3)

The code assigned by the Powder Diffraction File (JCPDS/ICDD).

[database]

**database.code\_depnum\_ccdc\_archive** (line)

database\_code\_depnum\_ccdc\_archive (cif\_core.dic 2.3)

Deposition numbers assigned by the Cambridge Crystallographic Data Centre (CCDC) to files containing structural information archived by the CCDC.

[database]

**database.code\_depnum\_ccdc\_fiz** (line)

database\_code\_depnum\_ccdc\_fiz (cif\_core.dic 2.3)

Deposition numbers assigned by the Fachinformationszentrum Karlsruhe (FIZ) to files containing structural information archived by the Cambridge Crystallographic Data Centre (CCDC).

[database]

**database.code\_depnum\_ccdc\_journal** (line)

database\_code\_depnum\_ccdc\_journal (cif\_core.dic 2.3)

Deposition numbers assigned by various journals to files containing structural information archived by the Cambridge Crystallographic Data Centre (CCDC).

[database]

**database.CSD\_history** (text)

database\_CSD\_history (cif\_core.dic 2.3)

A history of changes made by the Cambridge Crystallographic Data Centre and incorporated into the Cambridge Structural Database (CSD).

[database]

\* **database.entry\_id**

This data item is a pointer to entry.id in the ENTRY category.

**database.journal\_ASTM** (line)

database\_journal\_ASTM (cif\_core.dic 2.0.1)

The ASTM CODEN designator for a journal as given in the *Chemical Source List* maintained by the *Chemical Abstracts* Service.

[database]

**database.journal\_CSD** (line)

database\_journal\_CSD (cif\_core.dic 2.0.1)

The journal code used in the Cambridge Structural Database.

[database]

**DATABASE\_2**

Data items in the DATABASE\_2 category record details about the database identifiers of the data block. These data items are assigned by database managers and should only appear in a data block if they originate from that source. The name of this category, DATABASE\_2, arose because the category name DATABASE was already in use in the core CIF dictionary, but was used differently from the way it needed to be used in the mmCIF dictionary. Since CIF data names cannot be changed once they have been adopted, a new category had to be created.

Category group(s): **inclusive\_group**

**database\_group**

Category key(s): **database\_2.database\_id**

**database\_2.database\_code**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
database_2.database_id      'PDB'
database_2.database_code   '5HVP'
```

\* **database\_2.database\_code** (line)

The code assigned by the database identified in database\_2.database\_id.

Related items: database.code\_CAS (replaces),

database.code\_CSD (replaces),

database.code\_ICSD (replaces),

database.code\_MDF (replaces),

database.code\_NBS (replaces),

database.code\_PDF (replaces).

Examples: '1ABC', 'ABCDEF'.

[database\_2]

\* **database\_2.database\_id** (ucode)

An abbreviation that identifies the database.

Related items: database.code\_CAS (replaces),

database.code\_CSD (replaces),

database.code\_ICSD (replaces),

database.code\_MDF (replaces),

database.code\_NBS (replaces),

database.code\_PDF (replaces).

The data value must be one of the following:

CAS	<i>Chemical Abstracts</i>
CSD	Cambridge Structural Database (organic and metal-organic compounds)
ICSD	Inorganic Crystal Structure Database
MDF	Metals Data File (metal structures)
NDB	Nucleic Acid Database
NBS	NBS (NIST) Crystal Data Database (lattice parameters)
PDB	Protein Data Bank
PDF	Powder Diffraction File (JCPDS/ICDD)
RCSB	Research Collaboratory for Structural Bioinformatics
EBI	European Bioinformatics Institute

[database\_2]

**DATABASE\_PDB\_CAVEAT**

Data items in the DATABASE\_PDB\_CAVEAT category record details about features of the data block flagged as 'caveats' by the Protein Data Bank (PDB). These data items are included only for consistency with PDB format files. They should appear in a data block only if that data block was created by reformatting a PDB format file.

Category group(s): **inclusive\_group**

**database\_group**

**pdb\_group**

Category key(s): **database\_PDB\_caveat.id**

*Example 1 – hypothetical example.*

```
loop_
  _database_PDB_caveat.id
  _database_PDB_caveat.text
1
; THE CRYSTAL TRANSFORMATION IS IN ERROR BUT IS
;
2
; UNCORRECTABLE AT THIS TIME
;
```

\* **\_database\_PDB\_caveat.id** (int)  
A unique identifier for the PDB caveat record.

[database\_PDB\_caveat]

**\_database\_PDB\_caveat.text** (text)  
The full text of the PDB caveat record.

[database\_PDB\_caveat]

**DATABASE\_PDB\_MATRIX**

The DATABASE\_PDB\_MATRIX category provides placeholders for transformation matrices and vectors used by the Protein Data Bank (PDB). These data items are included only for consistency with older PDB format files. They should appear in a data block only if that data block was created by reformatting a PDB format file.

Category group(s): **inclusive\_group**  
**database\_group**  
**pdb\_group**

Category key(s): **\_database\_PDB\_matrix.entry\_id**

\* **\_database\_PDB\_matrix.entry\_id**  
This data item is a pointer to **\_entry.id** in the ENTRY category.

**\_database\_PDB\_matrix.origx[1][1]** (float)  
The [1][1] element of the PDB ORIGX matrix.  
Where no value is given, the assumed value is '1.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.origx[1][2]** (float)  
The [1][2] element of the PDB ORIGX matrix.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.origx[1][3]** (float)  
The [1][3] element of the PDB ORIGX matrix.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.origx[2][1]** (float)  
The [2][1] element of the PDB ORIGX matrix.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.origx[2][2]** (float)  
The [2][2] element of the PDB ORIGX matrix.  
Where no value is given, the assumed value is '1.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.origx[2][3]** (float)  
The [2][3] element of the PDB ORIGX matrix.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.origx[3][1]** (float)  
The [3][1] element of the PDB ORIGX matrix.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.origx[3][2]** (float)  
The [3][2] element of the PDB ORIGX matrix.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.origx[3][3]** (float)  
The [3][3] element of the PDB ORIGX matrix.  
Where no value is given, the assumed value is '1.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.origx\_vector[1]** (float)  
The [1] element of the PDB ORIGX vector.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.origx\_vector[2]** (float)  
The [2] element of the PDB ORIGX vector.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.origx\_vector[3]** (float)  
The [3] element of the PDB ORIGX vector.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.scale[1][1]** (float)  
The [1][1] element of the PDB SCALE matrix.  
Where no value is given, the assumed value is '1.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.scale[1][2]** (float)  
The [1][2] element of the PDB SCALE matrix.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.scale[1][3]** (float)  
The [1][3] element of the PDB SCALE matrix.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.scale[2][1]** (float)  
The [2][1] element of the PDB SCALE matrix.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.scale[2][2]** (float)  
The [2][2] element of the PDB SCALE matrix.  
Where no value is given, the assumed value is '1.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.scale[2][3]** (float)  
The [2][3] element of the PDB SCALE matrix.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.scale[3][1]** (float)  
The [3][1] element of the PDB SCALE matrix.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.scale[3][2]** (float)  
The [3][2] element of the PDB SCALE matrix.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.scale[3][3]** (float)  
The [3][3] element of the PDB SCALE matrix.  
Where no value is given, the assumed value is '1.0'. [database\_PDB\_matrix]

**\_database\_PDB\_matrix.scale\_vector[1]** (float)  
The [1] element of the PDB SCALE vector.  
Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

database\_PDB\_matrix.scale\_vector[2] (float)  
 The [2] element of the PDB SCALE vector.  
 Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

database\_PDB\_matrix.scale\_vector[3] (float)  
 The [3] element of the PDB SCALE vector.  
 Where no value is given, the assumed value is '0.0'. [database\_PDB\_matrix]

**DATABASE\_PDB\_REMARK**

Data items in the DATABASE\_PDB\_REMARK category record details about the data block as archived by the Protein Data Bank (PDB). Some data appearing in PDB REMARK records can be algorithmically extracted into the appropriate data items in the data block. These data items are included only for consistency with older PDB format files. They should appear in a data block only if that data block was created by reformatting a PDB format file.

Category group(s): inclusive\_group  
                   database\_group  
                   pdb\_group

Category key(s): database\_PDB\_remark.id

---

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```

loop_
  _database_PDB_remark.id
  _database_PDB_remark.text
3
; REFINEMENT. BY THE RESTRAINED LEAST-SQUARES PROCEDURE OF J.
KONNERT AND W. HENDRICKSON (PROGRAM *PROLSQ*). THE R
VALUE IS 0.176 FOR 12901 REFLECTIONS IN THE RESOLUTION
RANGE 8.0 TO 2.0 ANGSTROMS WITH I .GT. SIGMA(I).

RMS DEVIATIONS FROM IDEAL VALUES (THE VALUES OF
SIGMA, IN PARENTHESES, ARE THE INPUT ESTIMATED
STANDARD DEVIATIONS THAT DETERMINE THE RELATIVE
WEIGHTS OF THE CORRESPONDING RESTRAINTS)
DISTANCE RESTRAINTS (ANGSTROMS)
BOND DISTANCE                0.018 (0.020)
ANGLE DISTANCE                0.038 (0.030)
PLANAR 1-4 DISTANCE          0.043 (0.040)
PLANE RESTRAINT (ANGSTROMS)  0.015 (0.020)
CHIRAL-CENTER RESTRAINT (ANGSTROMS**3) 0.177 (0.150)
NON-BONDED CONTACT RESTRAINTS (ANGSTROMS)
SINGLE TORSION CONTACT        0.216 (0.500)
MULTIPLE TORSION CONTACT     0.207 (0.500)
POSSIBLE HYDROGEN BOND       0.245 (0.500)
CONFORMATIONAL TORSION ANGLE RESTRAINT (DEGREES)
PLANAR (OMEGA)               2.6 (3.0)
STAGGERED                    17.4 (15.0)
ORTHONORMAL                  18.1 (20.0)
;
4
; THE TWO CHAINS OF THE DIMERIC ENZYME HAS BEEN ASSIGNED THE
THE CHAIN INDICATORS *A* AND *B*.
;
# - - - data truncated for brevity - - -
        
```

\* database\_PDB\_remark.id (int)  
 A unique identifier for the PDB remark record.  
 [database\_PDB\_remark]

database\_PDB\_remark.text (text)  
 The full text of the PDB remark record.  
 [database\_PDB\_remark]

**DATABASE\_PDB\_REV**

Data items in the DATABASE\_PDB\_REV category record details about the history of the data block as archived by the Protein Data Bank (PDB). These data items are assigned by the PDB database managers and should only appear in a data block if they originate from that source.

Category group(s): inclusive\_group  
                   database\_group  
                   pdb\_group

Category key(s): database\_PDB\_rev.num

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```

loop_
  _database_PDB_rev.num
  _database_PDB_rev.author_name
  _database_PDB_rev.date
  _database_PDB_rev.date_original
  _database_PDB_rev.status
  _database_PDB_rev.mod_type
1 'Fitzgerald, Paula M.D' 1991-10-15 1990-04-30
'full release' 0
        
```

database\_PDB\_rev.author\_name (line)  
 The name of the person responsible for submitting this revision to the PDB. The family name(s) followed by a comma precedes the first name(s) or initial(s).  
 Examples: 'Bleary, Percival R.', 'O'Neil, F.K.', 'Van den Bossche, G.', 'Yang, D.-L.', 'Simonov, Yu.A.'. [database\_PDB\_rev]

database\_PDB\_rev.date (yyyy-mm-dd)  
 Date the PDB revision took place. Taken from the REVDAT record.  
 [database\_PDB\_rev]

database\_PDB\_rev.date\_original (yyyy-mm-dd)  
 Date the entry first entered the PDB database in the form yyyy-mm-dd. Taken from the PDB HEADER record.  
 Example: '1980-08-21'. [database\_PDB\_rev]

database\_PDB\_rev.mod\_type (int)  
 Taken from the REVDAT record. Refer to the Protein Data Bank format description at <http://www.rcsb.org/pdb/docs/format/pdbguide2.2/guide2.2.frame.html> for details.  
 The data value must be one of the following:

0	initial entry
1	all other types of modification
2	modifications to CONECT records
3	modifications affecting the coordinates or their transforms (CRYST1, ORIGX, SCALE, MTRIX, TVECT, ATOM, HETATM, SIGATM records)
4	layer 1 to layer 2 revision which may affect all record types
5	data uniformity processing

[database\_PDB\_rev]

\* database\_PDB\_rev.num (int)  
 The value of database\_PDB\_rev.num must uniquely and sequentially identify a record in the DATABASE\_PDB\_REV list. Note that this item must be a number and that modification numbers are assigned in increasing numerical order.

The following item(s) have an equivalent role in their respective categories:

database\_PDB\_rev.record.rev\_num. [database\_PDB\_rev]

database\_PDB\_rev.replaced\_by (line)  
 The PDB code for a subsequent PDB entry that replaced the PDB file corresponding to this data block.  
 [database\_PDB\_rev]



**`_diffrn.ambient_pressure_esd`** (float)  
 The standard uncertainty (estimated standard deviation) of `_diffrn.ambient_pressure`.  
 Related item: `_diffrn.ambient_pressure` (associated value). [diffrn]

**`_diffrn.ambient_pressure_gt`** (float)  
**`_diffrn_ambient_pressure_gt`** (cif.core.dic 2.3)  
 The mean hydrostatic pressure in kilopascals above which the intensities were measured. `_diffrn.ambient_pressure_gt` and `_diffrn.ambient_pressure_lt` allow a pressure range to be given. `_diffrn.ambient_pressure` should always be used in preference to these two items whenever possible.  
 The permitted range is [0.0, ∞).  
 Related item: `_diffrn.ambient_pressure` (alternate). [diffrn]

**`_diffrn.ambient_pressure_lt`** (float)  
**`_diffrn_ambient_pressure_lt`** (cif.core.dic 2.3)  
 The mean hydrostatic pressure in kilopascals below which the intensities were measured. `_diffrn.ambient_pressure_gt` and `_diffrn.ambient_pressure_lt` allow a pressure range to be given. `_diffrn.ambient_pressure` should always be used in preference to these two items whenever possible.  
 The permitted range is [0.0, ∞).  
 Related item: `_diffrn.ambient_pressure` (alternate). [diffrn]

**`_diffrn.ambient_temp`** (float, su)  
**`_diffrn_ambient_temperature`** (cif.core.dic 2.0.1)  
 The mean temperature in kelvins at which the intensities were measured.  
 The permitted range is [0.0, ∞).  
 Related item: `_diffrn.ambient_temp_esd` (associated esd). [diffrn]

**`_diffrn.ambient_temp_details`** (text)  
 A description of special aspects of temperature control during data collection. [diffrn]

**`_diffrn.ambient_temp_esd`** (float)  
 The standard uncertainty (estimated standard deviation) of `_diffrn.ambient_temp`.  
 Related item: `_diffrn.ambient_temp` (associated value). [diffrn]

**`_diffrn.ambient_temp_gt`** (float)  
**`_diffrn_ambient_temperature_gt`** (cif.core.dic 2.3)  
 The mean temperature in kelvins above which the intensities were measured. `_diffrn.ambient_temp_gt` and `_diffrn.ambient_temp_lt` allow a range of temperatures to be given. `_diffrn.ambient_temp` should always be used in preference to these two items whenever possible.  
 The permitted range is [0.0, ∞).  
 Related item: `_diffrn.ambient_temp` (alternate). [diffrn]

**`_diffrn.ambient_temp_lt`** (float)  
**`_diffrn_ambient_temperature_lt`** (cif.core.dic 2.3)  
 The mean temperature in kelvins below which the intensities were measured. `_diffrn.ambient_temp_gt` and `_diffrn.ambient_temp_lt` allow a range of temperatures to be given. `_diffrn.ambient_temp` should always be used in preference to these two items whenever possible.  
 The permitted range is [0.0, ∞).  
 Related item: `_diffrn.ambient_temp` (alternate). [diffrn]

\* **`_diffrn.crystal_id`**  
**`_diffrn_reflnt_crystal_id`** (cif.core.dic 2.0.1)  
 This data item is a pointer to `_exptl_crystal_id` in the EXPTL\_CRYSTAL category.

**`_diffrn.crystal_support`** (text)  
 The physical device used to support the crystal during data collection.  
 Examples: 'glass capillary', 'quartz capillary', 'fiber', 'metal loop'. [diffrn]

**`_diffrn.crystal_treatment`** (text)  
**`_diffrn_crystal_treatment`** (cif.core.dic 2.0.1)  
 Remarks about how the crystal was treated prior to intensity measurement. Particularly relevant when intensities were measured at low temperature.  
 Examples: 'equilibrated in hutch for 24 hours', 'flash frozen in liquid nitrogen', 'slow cooled with direct air stream'. [diffrn]

**`_diffrn.details`** (text)  
**`_diffrn_special_details`** (cif.core.dic 2.0.1)  
 Special details of the diffraction measurement process. Should include information about source instability, crystal motion, degradation and so on. [diffrn]

\* **`_diffrn.id`** (code)  
 This data item uniquely identifies a set of diffraction data.  
 The following item(s) have an equivalent role in their respective categories:  
`_diffrn_detector.diffrn_id`,  
`_diffrn_measurement.diffrn_id`,  
`_diffrn_orient_matrix.diffrn_id`,  
`_diffrn_orient_reflnt.diffrn_id`,  
`_diffrn_radiation.diffrn_id`,  
`_diffrn_reflnt.diffrn_id`,  
`_diffrn_reflnt.diffrn_id`,  
`_diffrn_source.diffrn_id`,  
`_diffrn_standard_reflnt.diffrn_id`,  
`_diffrn_standards.diffrn_id`. [diffrn]

#### DIFFRN\_ATTENUATOR

Data items in the DIFFRN\_ATTENUATOR category record details about the diffraction attenuator scales employed.

Category group(s): `inclusive_group`  
`diffrn_group`

Category key(s): `_diffrn_attenuator.code`

Example 2 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276-2277].

<code>_diffrn_attenuator.code</code>	1
<code>_diffrn_attenuator.scale</code>	16.976

\* **`_diffrn_attenuator.code`** (code)  
**`_diffrn_attenuator_code`** (cif.core.dic 2.0.1)  
 A code associated with a particular attenuator setting. This code is referenced by the `_diffrn_reflnt.attenuator_code` which is stored with the diffraction data. See `_diffrn_attenuator.scale`. [diffrn\_attenuator]

**`_diffrn_attenuator.material`** (text)  
**`_diffrn_attenuator_material`** (cif.core.dic 2.3)  
 Material from which the attenuator is made. [diffrn\_attenuator]

**\_diffrn\_attenuator.scale** (float)  
\_diffrn\_attenuator\_scale (cif\_core.dic 2.0.1)

The scale factor applied when an intensity measurement is reduced by an attenuator identified by \_diffrn\_attenuator.code. The measured intensity must be multiplied by this scale to convert it to the same scale as unattenuated intensities.

The permitted range is [1.0, ∞). [diffrn\_attenuator]

### DIFFRN\_DETECTOR

Data items in the DIFFRN\_DETECTOR category describe the detector used to measure the scattered radiation, including any analyser and post-sample collimation.

Category group(s): inclusive\_group

diffrn\_group

Category key(s): \_diffrn\_detector.diffrn\_id

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
_diffrn_detector.diffrn_id      'd1'
_diffrn_detector.detector      'multiwire'
_diffrn_detector.type          'Siemens'
```

**\_diffrn\_detector.area\_resol\_mean** (float)  
\_diffrn\_detector\_area\_resol\_mean (cif\_core.dic 2.3)

The resolution of an area detector, in pixels mm<sup>-1</sup>.

The permitted range is [0.0, ∞). [diffrn\_detector]

**\_diffrn\_detector.details** (text)  
\_diffrn\_detector\_details (cif\_core.dic 2.0.1)

A description of special aspects of the radiation detector.

[diffrn\_detector]

**\_diffrn\_detector.detector** (text)  
\_diffrn\_radiation\_detector (cifdic.c91 1.0)  
\_diffrn\_detector (cif\_core.dic 2.0)

The general class of the radiation detector.

Examples: 'photographic film', 'scintillation counter', 'CCD plate', 'BF-3~ counter'. [diffrn\_detector]

\* **\_diffrn\_detector.diffrn\_id**  
This data item is a pointer to \_diffrn\_id in the DIFFRN category.

**\_diffrn\_detector.dtime** (float)  
\_diffrn\_detector\_dtime (cif\_core.dic 2.3)

The deadtime in microseconds of the detector used to measure the diffraction intensities.

The permitted range is [0.0, ∞). [diffrn\_detector]

**\_diffrn\_detector.type** (text)  
\_diffrn\_detector\_type (cif\_core.dic 2.0.1)

The make, model or name of the detector device used.

[diffrn\_detector]

### DIFFRN\_MEASUREMENT

Data items in the DIFFRN\_MEASUREMENT category record details about the device used to orient and/or position the crystal during data measurement and the manner in which the diffraction data were measured.

Category group(s): inclusive\_group

diffrn\_group

Category key(s): \_diffrn\_measurement.diffrn\_id

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
_diffrn_measurement.diffrn_id  'd1'
_diffrn_measurement.device     '3-circle camera'
_diffrn_measurement.device_type 'Supper model x'
_diffrn_measurement.device_details 'none'
_diffrn_measurement.method     'omega scan'
_diffrn_measurement.details
; 440 frames, 0.20 degrees, 150 sec, detector distance 12 cm,
  detector angle 22.5 degrees
;
```

*Example 2 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].*

```
_diffrn_measurement.diffrn_id  's1'
_diffrn_measurement.device_type
  'Philips PW1100/20 diffractometer'
_diffrn_measurement.method     '\q/2\q'
```

**\_diffrn\_measurement.details** (text)  
\_diffrn\_measurement\_details (cif\_core.dic 2.0.1)

A description of special aspects of the intensity measurement.

Example:

```
; 440 frames, 0.20 degrees, 150 sec, detector
  distance 12 cm, detector angle 22.5 degrees
; [diffrn_measurement]
```

**\_diffrn\_measurement.device** (text)  
\_diffrn\_measurement\_device (cif\_core.dic 2.0.1)

The general class of goniometer or device used to support and orient the specimen.

Examples: '3-circle camera', '4-circle camera', 'kappa-geometry camera', 'oscillation camera', 'precession camera'. [diffrn\_measurement]

**\_diffrn\_measurement.device\_details** (text)  
\_diffrn\_measurement\_device\_details (cif\_core.dic 2.0.1)

A description of special aspects of the device used to measure the diffraction intensities.

Example:

```
; commercial goniometer modified locally to
  allow for 90\% \t arc
; [diffrn_measurement]
```

**\_diffrn\_measurement.device\_type** (text)  
\_diffrn\_measurement\_device\_type (cif\_core.dic 2.0.1)

The make, model or name of the measurement device (goniometer) used.

Examples: 'Supper model q', 'Huber model r', 'Enraf-Nonius model s', 'homemade'. [diffrn\_measurement]

\* **\_diffrn\_measurement.diffrn\_id**  
This data item is a pointer to \_diffrn\_id in the DIFFRN category.

**\_diffrn\_measurement.method** (text)  
\_diffrn\_measurement\_method (cif\_core.dic 2.0.1)

Method used to measure intensities.

Example: 'profile data from theta/2theta scans'. [diffrn\_measurement]

**diffrn\_measurement.specimen\_support** (text)  
diffrn\_measurement\_specimen\_support (cif\_core.dic 2.0.1)

The physical device used to support the crystal during data collection.

Examples: 'glass capillary', 'quartz capillary', 'fiber', 'metal loop'.  
 [diffrn\_measurement]

### DIFFRN\_ORIENT\_MATRIX

Data items in the DIFFRN\_ORIENT\_MATRIX category record details about the orientation matrix used in the measurement of the diffraction data.

Category group(s): inclusive\_group  
 diffrn\_group

Category key(s): diffrn\_orient\_matrix.diffrn\_id

Example 1 – based on CAD-4 diffractometer data obtained for Yb(S-C<sub>5</sub>H<sub>4</sub>N)<sub>2</sub>·(THF)<sub>4</sub>.

```

_diffrn_orient_matrix.diffrn_id  set1
_diffrn_orient_matrix.type
; reciprocal axis matrix, multiplies hkl vector to generate
  diffractometer xyz vector and diffractometer angles
;
_diffrn_orient_matrix.UB[1][1]  -0.071479
_diffrn_orient_matrix.UB[1][2]   0.020208
_diffrn_orient_matrix.UB[1][3]   0.039076
_diffrn_orient_matrix.UB[2][1]   0.035372
_diffrn_orient_matrix.UB[2][2]   0.056209
_diffrn_orient_matrix.UB[2][3]   0.078324
_diffrn_orient_matrix.UB[3][1]  -0.007470
_diffrn_orient_matrix.UB[3][2]   0.067854
_diffrn_orient_matrix.UB[3][3]  -0.017832

```

\* diffrn\_orient\_matrix.diffrn\_id

This data item is a pointer to diffrn\_id in the DIFFRN category.

**diffrn\_orient\_matrix.type** (text)  
diffrn\_orient\_matrix\_type (cif\_core.dic 2.0.1)

A description of the orientation matrix type and how it should be applied to define the orientation of the crystal precisely with respect to the diffractometer axes.

[diffrn\_orient\_matrix]

**diffrn\_orient\_matrix.UB[1][1]** (float)  
diffrn\_orient\_matrix\_UB\_11 (cif\_core.dic 2.0.1)

The [1][1] element of the 3 × 3 matrix that defines the dimensions of the reciprocal cell and its orientation with respect to the local diffractometer axes. See also diffrn\_orient\_matrix.type.

[diffrn\_orient\_matrix]

**diffrn\_orient\_matrix.UB[1][2]** (float)  
diffrn\_orient\_matrix\_UB\_12 (cif\_core.dic 2.0.1)

The [1][2] element of the 3 × 3 matrix that defines the dimensions of the reciprocal cell and its orientation with respect to the local diffractometer axes. See also diffrn\_orient\_matrix.type.

[diffrn\_orient\_matrix]

**diffrn\_orient\_matrix.UB[1][3]** (float)  
diffrn\_orient\_matrix\_UB\_13 (cif\_core.dic 2.0.1)

The [1][3] element of the 3 × 3 matrix that defines the dimensions of the reciprocal cell and its orientation with respect to the local diffractometer axes. See also diffrn\_orient\_matrix.type.

[diffrn\_orient\_matrix]

**diffrn\_orient\_matrix.UB[2][1]** (float)  
diffrn\_orient\_matrix\_UB\_21 (cif\_core.dic 2.0.1)

The [2][1] element of the 3 × 3 matrix that defines the dimensions of the reciprocal cell and its orientation with respect to the local diffractometer axes. See also diffrn\_orient\_matrix.type.

[diffrn\_orient\_matrix]

**diffrn\_orient\_matrix.UB[2][2]** (float)  
diffrn\_orient\_matrix\_UB\_22 (cif\_core.dic 2.0.1)

The [2][2] element of the 3 × 3 matrix that defines the dimensions of the reciprocal cell and its orientation with respect to the local diffractometer axes. See also diffrn\_orient\_matrix.type.

[diffrn\_orient\_matrix]

**diffrn\_orient\_matrix.UB[2][3]** (float)  
diffrn\_orient\_matrix\_UB\_23 (cif\_core.dic 2.0.1)

The [2][3] element of the 3 × 3 matrix that defines the dimensions of the reciprocal cell and its orientation with respect to the local diffractometer axes. See also diffrn\_orient\_matrix.type.

[diffrn\_orient\_matrix]

**diffrn\_orient\_matrix.UB[3][1]** (float)  
diffrn\_orient\_matrix\_UB\_31 (cif\_core.dic 2.0.1)

The [3][1] element of the 3 × 3 matrix that defines the dimensions of the reciprocal cell and its orientation with respect to the local diffractometer axes. See also diffrn\_orient\_matrix.type.

[diffrn\_orient\_matrix]

**diffrn\_orient\_matrix.UB[3][2]** (float)  
diffrn\_orient\_matrix\_UB\_32 (cif\_core.dic 2.0.1)

The [3][2] element of the 3 × 3 matrix that defines the dimensions of the reciprocal cell and its orientation with respect to the local diffractometer axes. See also diffrn\_orient\_matrix.type.

[diffrn\_orient\_matrix]

**diffrn\_orient\_matrix.UB[3][3]** (float)  
diffrn\_orient\_matrix\_UB\_33 (cif\_core.dic 2.0.1)

The [3][3] element of the 3 × 3 matrix that defines the dimensions of the reciprocal cell and its orientation with respect to the local diffractometer axes. See also diffrn\_orient\_matrix.type.

[diffrn\_orient\_matrix]

### DIFFRN\_ORIENT\_REFLN

Data items in the DIFFRN\_ORIENT\_REFLN category record details about the reflections that define the orientation matrix used in the measurement of the diffraction intensities.

Category group(s): inclusive\_group  
 diffrn\_group

Category key(s): diffrn\_orient\_refl.diffrn\_id  
diffrn\_orient\_refl.index\_h  
diffrn\_orient\_refl.index\_k  
diffrn\_orient\_refl.index\_l

Example 1 – based on CAD-4 diffractometer data obtained for Yb(S-C<sub>5</sub>H<sub>4</sub>N)<sub>2</sub>·(THF)<sub>4</sub>.

```

_diffrn_orient_refl.diffrn_id  myset1
_diffrn_orient_refl.index_h    2
_diffrn_orient_refl.index_k    0
_diffrn_orient_refl.index_l    2
_diffrn_orient_refl.angle_chi  -28.45
_diffrn_orient_refl.angle_kappa -11.32
_diffrn_orient_refl.angle_omega  5.33
_diffrn_orient_refl.angle_phi   101.78
_diffrn_orient_refl.angle_psi    0.00
_diffrn_orient_refl.angle_theta  10.66
# ... data abbreviated ...

```

**diffrn\_orient\_refl.angle\_chi** (float)

diffrn\_orient\_refl.angle\_chi(cif\_core.dic 2.0.1)

Diffraction angle  $\chi$  of a reflection used to define the orientation matrix in degrees. See diffrn\_orient\_matrix.UB and the Miller indices in the DIFFRN\_ORIENT\_REFLN category.

[diffrn\_orient\_refl]

**diffrn\_orient\_refl.angle\_kappa** (float)

diffrn\_orient\_refl.angle\_kappa(cif\_core.dic 2.0.1)

Diffraction angle  $\kappa$  of a reflection used to define the orientation matrix in degrees. See diffrn\_orient\_matrix.UB and the Miller indices in the DIFFRN\_ORIENT\_REFLN category.

[diffrn\_orient\_refl]

**diffrn\_orient\_refl.angle\_omega** (float)

diffrn\_orient\_refl.angle\_omega(cif\_core.dic 2.0.1)

Diffraction angle  $\omega$  of a reflection used to define the orientation matrix in degrees. See diffrn\_orient\_matrix.UB and the Miller indices in the DIFFRN\_ORIENT\_REFLN category.

[diffrn\_orient\_refl]

**diffrn\_orient\_refl.angle\_phi** (float)

diffrn\_orient\_refl.angle\_phi(cif\_core.dic 2.0.1)

Diffraction angle  $\varphi$  of a reflection used to define the orientation matrix in degrees. See diffrn\_orient\_matrix.UB and the Miller indices in the DIFFRN\_ORIENT\_REFLN category.

[diffrn\_orient\_refl]

**diffrn\_orient\_refl.angle\_psi** (float)

diffrn\_orient\_refl.angle\_psi(cif\_core.dic 2.0.1)

Diffraction angle  $\psi$  of a reflection used to define the orientation matrix in degrees. See diffrn\_orient\_matrix.UB and the Miller indices in the DIFFRN\_ORIENT\_REFLN category.

[diffrn\_orient\_refl]

**diffrn\_orient\_refl.angle\_theta** (float)

diffrn\_orient\_refl.angle\_theta(cif\_core.dic 2.0.1)

Diffraction angle  $\theta$  of a reflection used to define the orientation matrix in degrees. See diffrn\_orient\_matrix.UB and the Miller indices in the DIFFRN\_ORIENT\_REFLN category.

[diffrn\_orient\_refl]

\* **diffrn\_orient\_refl.diffrn\_id**

This data item is a pointer to diffrn\_id in the DIFFRN category.

\* **diffrn\_orient\_refl.index\_h** (int)

diffrn\_orient\_refl.index\_h(cif\_core.dic 2.0.1)

Miller index  $h$  of a reflection used to define the orientation matrix.

[diffrn\_orient\_refl]

\* **diffrn\_orient\_refl.index\_k** (int)

diffrn\_orient\_refl.index\_k(cif\_core.dic 2.0.1)

Miller index  $k$  of a reflection used to define the orientation matrix.

[diffrn\_orient\_refl]

\* **diffrn\_orient\_refl.index\_l** (int)

diffrn\_orient\_refl.index\_l(cif\_core.dic 2.0.1)

Miller index  $l$  of a reflection used to define the orientation matrix.

[diffrn\_orient\_refl]

## DIFFRN\_RADIATION

Data items in the DIFFRN\_RADIATION category describe the radiation used in measuring the diffraction intensities, its collimation and monochromatization before the sample. Post-sample treatment of the beam is described by data items in the DIFFRN\_DETECTOR category.

Category group(s): inclusive\_group

diffrn\_group

Category key(s): diffrn\_radiation.diffrn\_id

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
diffrn_radiation.diffrn_id          'set1'
diffrn_radiation.collimation      '0.3 mm double pinhole'
diffrn_radiation.monochromator    'graphite'
diffrn_radiation.type             'Cu K\alpha'
diffrn_radiation.wavelength_id    1
```

Example 2 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].

```
diffrn_radiation.wavelength_id    1
diffrn_radiation.type             'Cu K\alpha'
diffrn_radiation.monochromator    'graphite'
```

**diffrn\_radiation.collimation** (text)

diffrn\_radiation.collimation(cif\_core.dic 2.0.1)

The collimation or focusing applied to the radiation.

Examples: '0.3 mm double-pinhole', '0.5 mm', 'focusing mirrors'.

[diffrn\_radiation]

\* **diffrn\_radiation.diffrn\_id**

This data item is a pointer to diffrn\_id in the DIFFRN category.

**diffrn\_radiation.filter\_edge** (float)

diffrn\_radiation.filter\_edge(cif\_core.dic 2.0.1)

Absorption edge in ångströms of the radiation filter used.

The permitted range is [0.0, ∞).

[diffrn\_radiation]

**diffrn\_radiation.inhomogeneity** (float)

diffrn\_radiation.inhomogeneity(cif\_core.dic 2.0.1)

Half-width in millimetres of the incident beam in the direction perpendicular to the diffraction plane.

The permitted range is [0.0, ∞).

[diffrn\_radiation]

**diffrn\_radiation.monochromator** (text)

diffrn\_radiation.monochromator(cif\_core.dic 2.0.1)

The method used to obtain monochromatic radiation. If a monochromator crystal is used, the material and the indices of the Bragg reflection are specified.

Examples: 'Zr filter', 'Ge 220', 'none', 'equatorial mounted graphite'.

[diffrn\_radiation]

**diffrn\_radiation.polarisn\_norm** (float)

diffrn\_radiation.polarisn\_norm(cif\_core.dic 2.0.1)

The angle in degrees, as viewed from the specimen, between the perpendicular component of the polarization and the diffraction plane. See diffrn\_radiation.polarisn\_ratio.

The permitted range is [0.0, ∞).

[diffrn\_radiation]

**diffrn\_radiation.polarisn\_ratio** (float)

diffrn\_radiation.polarisn\_ratio(cif\_core.dic 2.0.1)

Polarization ratio of the diffraction beam incident on the crystal. This is the ratio of the perpendicularly polarized to the parallel-polarized component of the radiation. The perpendicular component forms an angle of diffrn\_radiation.polarisn\_norm to the normal to the diffraction plane of the sample (*i.e.* the plane containing the incident and reflected beams).

The permitted range is [0.0, ∞).

[diffrn\_radiation]

**diffrn\_radiation.probe** (line)

diffrn\_radiation\_probe (cif\_core.dic 2.0.1)

The nature of the radiation used (*i.e.* the name of the subatomic particle or the region of the electromagnetic spectrum). It is strongly recommended that this information be given, so that the probe radiation can be simply determined.

The data value must be one of the following:

x-ray  
neutron  
electron  
gamma

[diffrn\_radiation]

**diffrn\_radiation.type** (line)

diffrn\_radiation\_type (cif\_core.dic 2.0.1)

The nature of the radiation. This is typically a description of the X-ray wavelength in Siegbahn notation.

Examples: 'CuK\alpha', 'Cu K\alpha~1~', 'Cu K-L~2,3~', 'white-beam'.

[diffrn\_radiation]

\* **diffrn\_radiation.wavelength\_id**

This data item is a pointer to diffrn\_radiation\_wavelength\_id in the DIFFRN\_RADIATION\_WAVELENGTH category.

**diffrn\_radiation.xray\_symbol** (line)

diffrn\_radiation\_xray\_symbol (cif\_core.dic 2.0.1)

The IUPAC symbol for the X-ray wavelength for the probe radiation.

The data value must be one of the following:

K-L~3~                     $K\alpha_1$  in older Siegbahn notation  
K-L~2~                     $K\alpha_2$  in older Siegbahn notation  
K-M~3~                     $K\beta$  in older Siegbahn notation  
K-L~2,3~                 use where  $K-L_3$  and  $K-L_2$  are not resolved

[diffrn\_radiation]

### DIFFRN\_RADIATION\_WAVELENGTH

Data items in the DIFFRN\_RADIATION\_WAVELENGTH category describe the wavelength of the radiation used to measure the diffraction intensities. Items may be looped to identify and assign weights to distinct components of a polychromatic beam.

Category group(s): inclusive\_group  
diffrn\_group

Category key(s): diffrn\_radiation\_wavelength\_id

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

<u>diffrn_radiation_wavelength_id</u>	1
<u>diffrn_radiation_wavelength_wavelength</u>	1.54
<u>diffrn_radiation_wavelength_wt</u>	1.0

\* **diffrn\_radiation\_wavelength\_id** (code)

diffrn\_radiation\_wavelength\_id (cif\_core.dic 2.0.1)

The code identifying each value of diffrn\_radiation\_wavelength\_wavelength. Items in the DIFFRN\_RADIATION\_WAVELENGTH category are looped when multiple wavelengths are used. This code is used to link with the DIFFRN\_REFLN category. The diffrn\_refl.wavelength\_id codes must match one of the codes defined in this category.

The following item(s) have an equivalent role in their respective categories:

diffrn\_radiation\_wavelength\_id,  
diffrn\_refl.wavelength\_id,  
refl.wavelength\_id.

Examples: 'x1', 'x2', 'neut'.

[diffrn\_radiation\_wavelength]

\* **diffrn\_radiation\_wavelength.wavelength** (float)

diffrn\_radiation\_wavelength (cif\_core.dic 2.0.1)

The radiation wavelength in ångströms.

The permitted range is [0.0, ∞). [diffrn\_radiation\_wavelength]

**diffrn\_radiation\_wavelength.wt** (float)

diffrn\_radiation\_wavelength\_wt (cif\_core.dic 2.0.1)

The relative weight of a wavelength identified by the code diffrn\_radiation\_wavelength\_id in the list of wavelengths.

The permitted range is [0.0, 1.0]. Where no value is given, the assumed value is '1.0'.

[diffrn\_radiation\_wavelength]

### DIFFRN\_REFLN

Data items in the DIFFRN\_REFLN category record details about the intensities in the diffraction data set identified by diffrn\_refl.diffrn\_id. The DIFFRN\_REFLN data items refer to individual intensity measurements and must be included in looped lists. The DIFFRN\_REFLNS data items specify the parameters that apply to all intensity measurements in the particular diffraction data set identified by diffrn\_reflns.diffrn\_id.

Category group(s): inclusive\_group  
diffrn\_group

Category key(s): diffrn\_refl.diffrn\_id  
diffrn\_refl.id

Example 1 – based on CAD-4 diffractometer data obtained for  $\text{Yb}(\text{S-C}_5\text{H}_4\text{N})_2\text{-}(\text{THF})_4$  for data set 'set1' reflection 1102.

<u>diffrn_refl.diffrn_id</u>	set1
<u>diffrn_refl.id</u>	1102
<u>diffrn_refl.wavelength_id</u>	Culfixed
<u>diffrn_refl.angle_chi</u>	32.21
<u>diffrn_refl.angle_kappa</u>	20.12
<u>diffrn_refl.angle_omega</u>	11.54
<u>diffrn_refl.angle_phi</u>	176.02
<u>diffrn_refl.angle_psi</u>	0.00
<u>diffrn_refl.angle_theta</u>	23.08
<u>diffrn_refl.attenuator_code</u>	'Ni.005'
<u>diffrn_refl.counts_bg_1</u>	22
<u>diffrn_refl.counts_bg_2</u>	25
<u>diffrn_refl.counts_net</u>	3450
<u>diffrn_refl.counts_peak</u>	321
<u>diffrn_refl.counts_total</u>	3499
<u>diffrn_refl.detect_slit_horiz</u>	0.04
<u>diffrn_refl.detect_slit_vert</u>	0.02
<u>diffrn_refl.elapsed_time</u>	1.00
<u>diffrn_refl.index_h</u>	4
<u>diffrn_refl.index_k</u>	0
<u>diffrn_refl.index_l</u>	2
<u>diffrn_refl.intensity_net</u>	202.56
<u>diffrn_refl.intensity_sigma</u>	2.18
<u>diffrn_refl.scale_group_code</u>	A24
<u>diffrn_refl.scan_mode</u>	om
<u>diffrn_refl.scan_mode_backgd</u>	mo
<u>diffrn_refl.scan_rate</u>	1.2
<u>diffrn_refl.scan_time_backgd</u>	900.00
<u>diffrn_refl.scan_width</u>	1.0
<u>diffrn_refl.sint_over_lambda</u>	0.25426
<u>diffrn_refl.standard_code</u>	1
<u>diffrn_refl.wavelength</u>	1.54184

**diffrn\_refl.angle\_chi** (float)

diffrn\_refl\_angle\_chi (cif\_core.dic 2.0.1)

The diffractometer angle  $\chi$  of a reflection in degrees. This angle corresponds to the specified orientation matrix and the original measured cell before any subsequent cell transformations.

[diffrn\_refl]

**diffrn\_refl.angle\_kappa** (float)

diffrn\_refl\_angle\_kappa (cif\_core.dic 2.0.1)

The diffractometer angle  $\kappa$  of a reflection in degrees. This angle corresponds to the specified orientation matrix and the original measured cell before any subsequent cell transformations.

[diffrn\_refl]

**`_diffrn_refl.angle_omega`** (float)  
`_diffrn_refl.angle_omega` (cif\_core.dic 2.0.1)  
 The diffractometer angle  $\omega$  of a reflection in degrees. This angle corresponds to the specified orientation matrix and the original measured cell before any subsequent cell transformations.  
 [diffrn\_refl]

**`_diffrn_refl.angle_phi`** (float)  
`_diffrn_refl.angle_phi` (cif\_core.dic 2.0.1)  
 The diffractometer angle  $\varphi$  of a reflection in degrees. This angle corresponds to the specified orientation matrix and the original measured cell before any subsequent cell transformations.  
 [diffrn\_refl]

**`_diffrn_refl.angle_psi`** (float)  
`_diffrn_refl.angle_psi` (cif\_core.dic 2.0.1)  
 The diffractometer angle  $\psi$  of a reflection in degrees. This angle corresponds to the specified orientation matrix and the original measured cell before any subsequent cell transformations.  
 [diffrn\_refl]

**`_diffrn_refl.angle_theta`** (float)  
`_diffrn_refl.angle_theta` (cif\_core.dic 2.0.1)  
 The diffractometer angle  $\theta$  of a reflection in degrees. This angle corresponds to the specified orientation matrix and the original measured cell before any subsequent cell transformations.  
 [diffrn\_refl]

**`_diffrn_refl.attenuator_code`**  
`_diffrn_refl.attenuator_code` (cif\_core.dic 2.0.1)  
 The code identifying the attenuator setting for this reflection. This code must match one of the `_diffrn_attenuator.code` values.

**`_diffrn_refl.class_code`** (code)  
`_diffrn_refl.class_code` (cif\_core.dic 2.3)  
 The code identifying the class to which this reflection has been assigned. This code must match a value of `_diffrn_reflns.class_code`. Reflections may be grouped into classes for a variety of purposes. For example, for modulated structures each reflection class may be defined by the number  $m = \sum |m_i|$ , where the  $m_i$  are the integer coefficients that, in addition to  $h, k, l$ , index the corresponding diffraction vector in the basis defined for the reciprocal lattice.  
 [diffrn\_refl]

**`_diffrn_refl.counts_bg_1`** (int)  
`_diffrn_refl.counts_bg_1` (cif\_core.dic 2.0.1)  
 The diffractometer counts for the measurement of the background before the peak.  
 The permitted range is  $[0, \infty)$ .  
 [diffrn\_refl]

**`_diffrn_refl.counts_bg_2`** (int)  
`_diffrn_refl.counts_bg_2` (cif\_core.dic 2.0.1)  
 The diffractometer counts for the measurement of the background after the peak.  
 The permitted range is  $[0, \infty)$ .  
 [diffrn\_refl]

**`_diffrn_refl.counts_net`** (int)  
`_diffrn_refl.counts_net` (cif\_core.dic 2.0.1)  
 The diffractometer counts for the measurement of net counts after background removal.  
 The permitted range is  $[0, \infty)$ .  
 [diffrn\_refl]

**`_diffrn_refl.counts_peak`** (int)  
`_diffrn_refl.counts_peak` (cif\_core.dic 2.0.1)  
 The diffractometer counts for the measurement of counts for the peak scan or position.  
 The permitted range is  $[0, \infty)$ .  
 [diffrn\_refl]

**`_diffrn_refl.counts_total`** (int)  
`_diffrn_refl.counts_total` (cif\_core.dic 2.0.1)  
 The diffractometer counts for the measurement of total counts (background plus peak).  
 The permitted range is  $[0, \infty)$ .  
 [diffrn\_refl]

**`_diffrn_refl.detect_slit_horiz`** (float)  
`_diffrn_refl.detect_slit_horiz` (cif\_core.dic 2.0.1)  
 Total slit aperture in degrees in the diffraction plane.  
 The permitted range is  $[0.0, 90.0]$ .  
 [diffrn\_refl]

**`_diffrn_refl.detect_slit_vert`** (float)  
`_diffrn_refl.detect_slit_vert` (cif\_core.dic 2.0.1)  
 Total slit aperture in degrees perpendicular to the diffraction plane.  
 The permitted range is  $[0.0, 90.0]$ .  
 [diffrn\_refl]

\* **`_diffrn_refl.diffrn_id`**  
 This data item is a pointer to `_diffrn.id` in the DIFFRN category.

**`_diffrn_refl.elapsed_time`** (float)  
`_diffrn_refl.elapsed_time` (cif\_core.dic 2.0.1)  
 Elapsed time in minutes from the start of the diffraction experiment to the measurement of this intensity.  
 The permitted range is  $[0.0, \infty)$ .  
 [diffrn\_refl]

\* **`_diffrn_refl.id`** (code)  
 The value of `_diffrn_refl.id` must uniquely identify the reflection in the data set identified by the item `_diffrn_refl.diffrn_id`. Note that this item need not be a number; it can be any unique identifier.  
 [diffrn\_refl]

\* **`_diffrn_refl.index_h`** (int)  
`_diffrn_refl.index_h` (cif\_core.dic 2.0.1)  
 Miller index  $h$  of a reflection. The values of the Miller indices in the DIFFRN\_REFLN category need not match the values of the Miller indices in the REFLN category if a transformation of the original measured cell has taken place. Details of the cell transformation are given in `_diffrn_reflns.reduction_process`. See also `_diffrn_reflns.transf_matrix[][]`.  
 [diffrn\_refl]

\* **`_diffrn_refl.index_k`** (int)  
`_diffrn_refl.index_k` (cif\_core.dic 2.0.1)  
 Miller index  $k$  of a reflection. The values of the Miller indices in the DIFFRN\_REFLN category need not match the values of the Miller indices in the REFLN category if a transformation of the original measured cell has taken place. Details of the cell transformation are given in `_diffrn_reflns.reduction_process`. See also `_diffrn_reflns.transf_matrix[][]`.  
 [diffrn\_refl]

\* **`_diffrn_refl.index_l`** (int)  
`_diffrn_refl.index_l` (cif\_core.dic 2.0.1)  
 Miller index  $l$  of a reflection. The values of the Miller indices in the DIFFRN\_REFLN category need not match the values of the Miller indices in the REFLN category if a transformation of the original measured cell has taken place. Details of the cell transformation are given in `_diffrn_reflns.reduction_process`. See also `_diffrn_reflns.transf_matrix[][]`.  
 [diffrn\_refl]

**`_diffrn_refl.intensity_net`** (float)  
`_diffrn_refl.intensity_net` (cif.core.dic 2.0.1)  
 Net intensity calculated from the diffraction counts after the attenuator and standard scales have been applied.  
 The permitted range is [0, ∞). [diffrn\_refl]

**`_diffrn_refl.intensity_sigma`** (float)  
`_diffrn_refl.intensity_sigma` (cif.core.dic 2.0.1)  
 Standard uncertainty (estimated standard deviation) of the intensity calculated from the diffraction counts after the attenuator and standard scales have been applied.  
 The permitted range is [0, ∞). [diffrn\_refl]

**`_diffrn_refl.intensity_u`** (float)  
`_diffrn_refl.intensity_u` (cif.core.dic 2.3)  
 Standard uncertainty of the net intensity calculated from the diffraction counts after the attenuator and standard scales have been applied.  
 The permitted range is [0.0, ∞).  
 Related item: `_diffrn_refl.intensity_sigma` (alternate). [diffrn\_refl]

\* **`_diffrn_refl.scale_group_code`**  
`_diffrn_refl.scale_group_code` (cif.core.dic 2.0.1)  
 The code identifying the scale applying to this reflection. This data item is a pointer to `_diffrn_scale_group.code` in the DIFFRN\_SCALE\_GROUP category.

**`_diffrn_refl.scan_mode`** (ucode)  
`_diffrn_refl.scan_mode` (cif.core.dic 2.0.1)  
 The code identifying the mode of scanning for measurements using a diffractometer. See `_diffrn_refl.scan_width` and `_diffrn_refl.scan_mode_backgd`.  
 The data value must be one of the following:  
 om ω scan  
 ot ω/2θ scan  
 q Q scans (arbitrary reciprocal directions)  
 [diffrn\_refl]

**`_diffrn_refl.scan_mode_backgd`** (ucode)  
`_diffrn_refl.scan_mode_backgd` (cif.core.dic 2.0.1)  
 The code identifying the mode of scanning a reflection to measure the background intensity.  
 The data value must be one of the following:  
 st stationary counter background  
 mo moving counter background  
 [diffrn\_refl]

**`_diffrn_refl.scan_rate`** (float)  
`_diffrn_refl.scan_rate` (cif.core.dic 2.0.1)  
 The rate of scanning a reflection in degrees per minute to measure the intensity.  
 [diffrn\_refl]

**`_diffrn_refl.scan_time_backgd`** (float)  
`_diffrn_refl.scan_time_backgd` (cif.core.dic 2.0.1)  
 The time spent measuring each background in seconds.  
 [diffrn\_refl]

**`_diffrn_refl.scan_width`** (float)  
`_diffrn_refl.scan_width` (cif.core.dic 2.0.1)  
 The scan width in degrees of the scan mode defined by the code `_diffrn_refl.scan_mode`.  
 The permitted range is [0.0, 90.0]. [diffrn\_refl]

**`_diffrn_refl.sint_over_lambda`** (float)  
`_diffrn_refl.sint/lambda` (cif.core.dic 2.0.1)  
 The  $(\sin \theta)/\lambda$  value in reciprocal ångströms for this reflection.  
 The permitted range is [0.0, ∞). [diffrn\_refl]

\* **`_diffrn_refl.standard_code`**  
`_diffrn_refl.standard_code` (cif.core.dic 2.0.1)  
 The code identifying that this reflection was measured as a standard intensity. This data item is a pointer to `_diffrn_standard_refl.code` in the DIFFRN\_STANDARD\_REFLN category.

**`_diffrn_refl.wavelength`** (float)  
`_diffrn_refl.wavelength` (cif.core.dic 2.0.1)  
 The mean wavelength in ångströms of the radiation used to measure the intensity of this reflection. This is an important parameter for data collected using energy-dispersive detectors or the Laue method.  
 The permitted range is [0.0, ∞). [diffrn\_refl]

\* **`_diffrn_refl.wavelength_id`**  
`_diffrn_refl.wavelength_id` (cif.core.dic 2.0.1)  
 This data item is a pointer to `_diffrn_radiation.wavelength_id` in the DIFFRN\_RADIATION category.

### DIFFRN\_REFLNS

Data items in the DIFFRN\_REFLNS category record details about the set of intensities measured in the diffraction experiment. The DIFFRN\_REFLN data items refer to individual intensity measurements and must be included in looped lists. The DIFFRN\_REFLNS data items specify the parameters that apply to all intensity measurements in a diffraction data set.

Category group(s): `inclusive_group`  
`diffrn_group`  
 Category key(s): `_diffrn_reflns.diffrn_id`

**`_diffrn_reflns.av_R_equivalents`** (float)  
`_diffrn_reflns.av_R_equivalents` (cif.core.dic 2.0.1)  
 The residual  $[\sum \text{av}|\Delta(I)| / \sum |\text{av}(I)|]$  for symmetry-equivalent reflections used to calculate the average intensity  $\text{av}(I)$ . The  $\text{av}|\Delta(I)|$  term is the average absolute difference between  $\text{av}(I)$  and the individual symmetry-equivalent intensities.  
 The permitted range is [0.0, ∞). [diffrn\_reflns]

**`_diffrn_reflns.av_sigmaI_over_netI`** (float)  
`_diffrn_reflns.av_sigmaI/netI` (cif.core.dic 2.0.1)  
 Measure  $[\sum |\sigma(\text{net } I)| / \sum |\text{net } I|]$  for all measured reflections.  
 The permitted range is [0.0, ∞). [diffrn\_reflns]

**`_diffrn_reflns.av_unetI/netI`** (float)  
`_diffrn_reflns.av_unetI/netI` (cif.core.dic 2.3)  
 Measure  $[\sum |u(\text{net } I)| / \sum |\text{net } I|]$  for all measured reflections.  
 The permitted range is [0.0, ∞). [diffrn\_reflns]

\* **`_diffrn_reflns.diffrn_id`**  
 This data item is a pointer to `_diffrn.id` in the DIFFRN category.

**`_diffrn_reflns.limit_h_max`** (int)  
`_diffrn_reflns.limit_h_max` (cif.core.dic 2.0.1)  
 The maximum value of the Miller index  $h$  for the reflection data specified by `_diffrn_refl.index_h`. [diffrn\_reflns]

<p><b><code>_diffrn_reflns.limit_h_min</code></b> (int)  <code>_diffrn_reflns_limit_h_min</code> (cif_core.dic 2.0.1)                      The minimum value of the Miller index <math>h</math> for the reflection data specified by <code>_diffrn_refl_index_h</code>.</p>	<p><b><code>_diffrn_reflns.transf_matrix[1][2]</code></b> (float)  <code>_diffrn_reflns_transf_matrix_12</code> (cif_core.dic 2.0.1)                      The [1][2] element of the <math>3 \times 3</math> matrix used to transform Miller indices in the DIFFRN_REFLN category into the Miller indices in the REFLN category.</p>
<code>[diffrn_reflns]</code>	<code>[diffrn_reflns]</code>
<p><b><code>_diffrn_reflns.limit_k_max</code></b> (int)  <code>_diffrn_reflns_limit_k_max</code> (cif_core.dic 2.0.1)                      The maximum value of the Miller index <math>k</math> for the reflection data specified by <code>_diffrn_refl_index_k</code>.</p>	<p><b><code>_diffrn_reflns.transf_matrix[1][3]</code></b> (float)  <code>_diffrn_reflns_transf_matrix_13</code> (cif_core.dic 2.0.1)                      The [1][3] element of the <math>3 \times 3</math> matrix used to transform Miller indices in the DIFFRN_REFLN category into the Miller indices in the REFLN category.</p>
<code>[diffrn_reflns]</code>	<code>[diffrn_reflns]</code>
<p><b><code>_diffrn_reflns.limit_k_min</code></b> (int)  <code>_diffrn_reflns_limit_k_min</code> (cif_core.dic 2.0.1)                      The minimum value of the Miller index <math>k</math> for the reflection data specified by <code>_diffrn_refl_index_k</code>.</p>	<p><b><code>_diffrn_reflns.transf_matrix[2][1]</code></b> (float)  <code>_diffrn_reflns_transf_matrix_21</code> (cif_core.dic 2.0.1)                      The [2][1] element of the <math>3 \times 3</math> matrix used to transform Miller indices in the DIFFRN_REFLN category into the Miller indices in the REFLN category.</p>
<code>[diffrn_reflns]</code>	<code>[diffrn_reflns]</code>
<p><b><code>_diffrn_reflns.limit_l_max</code></b> (int)  <code>_diffrn_reflns_limit_l_max</code> (cif_core.dic 2.0.1)                      The maximum value of the Miller index <math>l</math> for the reflection data specified by <code>_diffrn_refl_index_l</code>.</p>	<p><b><code>_diffrn_reflns.transf_matrix[2][2]</code></b> (float)  <code>_diffrn_reflns_transf_matrix_22</code> (cif_core.dic 2.0.1)                      The [2][2] element of the <math>3 \times 3</math> matrix used to transform Miller indices in the DIFFRN_REFLN category into the Miller indices in the REFLN category.</p>
<code>[diffrn_reflns]</code>	<code>[diffrn_reflns]</code>
<p><b><code>_diffrn_reflns.limit_l_min</code></b> (int)  <code>_diffrn_reflns_limit_l_min</code> (cif_core.dic 2.0.1)                      The minimum value of the Miller index <math>l</math> for the reflection data specified by <code>_diffrn_refl_index_l</code>.</p>	<p><b><code>_diffrn_reflns.transf_matrix[2][3]</code></b> (float)  <code>_diffrn_reflns_transf_matrix_23</code> (cif_core.dic 2.0.1)                      The [2][3] element of the <math>3 \times 3</math> matrix used to transform Miller indices in the DIFFRN_REFLN category into the Miller indices in the REFLN category.</p>
<code>[diffrn_reflns]</code>	<code>[diffrn_reflns]</code>
<p><b><code>_diffrn_reflns.number</code></b> (int)  <code>_diffrn_reflns_number</code> (cif_core.dic 2.0.1)                      The total number of measured intensities, excluding reflections that are classified as systematically absent.                      The permitted range is <math>[0, \infty)</math>.</p>	<p><b><code>_diffrn_reflns.transf_matrix[3][1]</code></b> (float)  <code>_diffrn_reflns_transf_matrix_31</code> (cif_core.dic 2.0.1)                      The [3][1] element of the <math>3 \times 3</math> matrix used to transform Miller indices in the DIFFRN_REFLN category into the Miller indices in the REFLN category.</p>
<code>[diffrn_reflns]</code>	<code>[diffrn_reflns]</code>
<p><b><code>_diffrn_reflns.reduction_process</code></b> (text)  <code>_diffrn_reflns_reduction_process</code> (cif_core.dic 2.0.1)                      A description of the process used to reduce the intensity data into structure-factor magnitudes.                      Example: 'data averaged using Fisher test'.</p>	<p><b><code>_diffrn_reflns.transf_matrix[3][2]</code></b> (float)  <code>_diffrn_reflns_transf_matrix_32</code> (cif_core.dic 2.0.1)                      The [3][2] element of the <math>3 \times 3</math> matrix used to transform Miller indices in the DIFFRN_REFLN category into the Miller indices in the REFLN category.</p>
<code>[diffrn_reflns]</code>	<code>[diffrn_reflns]</code>
<p><b><code>_diffrn_reflns.theta_max</code></b> (float)  <code>_diffrn_reflns_theta_max</code> (cif_core.dic 2.0.1)                      Maximum <math>\theta</math> angle in degrees for the measured diffraction intensities.                      The permitted range is <math>[0.0, 90.0]</math>.</p>	<p><b><code>_diffrn_reflns.transf_matrix[3][3]</code></b> (float)  <code>_diffrn_reflns_transf_matrix_33</code> (cif_core.dic 2.0.1)                      The [3][3] element of the <math>3 \times 3</math> matrix used to transform Miller indices in the DIFFRN_REFLN category into the Miller indices in the REFLN category.</p>
<code>[diffrn_reflns]</code>	<code>[diffrn_reflns]</code>
<p><b><code>_diffrn_reflns.theta_min</code></b> (float)  <code>_diffrn_reflns_theta_min</code> (cif_core.dic 2.0.1)                      Minimum <math>\theta</math> angle in degrees for the measured diffraction intensities.                      The permitted range is <math>[0.0, 90.0]</math>.</p>	<p><b><code>_diffrn_reflns.transf_matrix[1][1]</code></b> (float)  <code>_diffrn_reflns_transf_matrix_11</code> (cif_core.dic 2.0.1)                      The [1][1] element of the <math>3 \times 3</math> matrix used to transform Miller indices in the DIFFRN_REFLN category into the Miller indices in the REFLN category.</p>
<code>[diffrn_reflns]</code>	<code>[diffrn_reflns]</code>
<p><b><code>_diffrn_reflns.transf_matrix[1][1]</code></b> (float)  <code>_diffrn_reflns_transf_matrix_11</code> (cif_core.dic 2.0.1)                      The [1][1] element of the <math>3 \times 3</math> matrix used to transform Miller indices in the DIFFRN_REFLN category into the Miller indices in the REFLN category.</p>	<p><b><code>_diffrn_reflns.transf_matrix[3][3]</code></b> (float)  <code>_diffrn_reflns_transf_matrix_33</code> (cif_core.dic 2.0.1)                      The [3][3] element of the <math>3 \times 3</math> matrix used to transform Miller indices in the DIFFRN_REFLN category into the Miller indices in the REFLN category.</p>
<code>[diffrn_reflns]</code>	<code>[diffrn_reflns]</code>

**DIFFRN\_REFLNS\_CLASS**

Data items in the DIFFRN\_REFLNS\_CLASS category record details about the classes of reflections measured in the diffraction experiment.

Category key(s): `_diffrn_reflns_class.code`

*Example 1 – example corresponding to the one-dimensional incommensurately modulated structure of  $K_2SeO_4$ . Each reflection class is defined by the number  $m = \sum |m_i|$ , where the  $m_i$  are the integer coefficients that, in addition to  $h, k, l$ , index the corresponding diffraction vector in the basis defined for the reciprocal lattice.*

```
loop_
  _diffrn_reflns_class.number
  _diffrn_reflns_class.d_res_high
  _diffrn_reflns_class.d_res_low
  _diffrn_reflns_class.av_R_eq
  _diffrn_reflns_class.code
  _diffrn_reflns_class.description
1580 0.551 6.136 0.015 'Main' 'm=0; main reflections'
1045 0.551 6.136 0.010 'Sat1' 'm=1; first-order satellites'
```

`_diffrn_reflns_class.av_R_eq` (float)  
`_diffrn_reflns_class.av_R_eq` (cif\_core.dic 2.3)

For each reflection class, the residual  $[\sum \text{av}|\Delta(I)| / \sum |\text{av}(I)|]$  for symmetry-equivalent reflections used to calculate the average intensity  $\text{av}(I)$ . The  $\text{av}|\Delta(I)|$  term is the average absolute difference between  $\text{av}(I)$  and the individual intensities.

The permitted range is  $[0.0, \infty)$ . [diffrn\_reflns\_class]

`_diffrn_reflns_class.av_sgI/I` (float)  
`_diffrn_reflns_class.av_sgI/I` (cif\_core.dic 2.3)

Measure  $[\sum |\sigma(\text{net } I)| / \sum |\text{net } I|]$  for all measured intensities in a reflection class.

The permitted range is  $[0.0, \infty)$ .

Related item: `_diffrn_reflns_class.av_uI/I` (replaces).

[diffrn\_reflns\_class]

`_diffrn_reflns_class.av_uI/I` (float)  
`_diffrn_reflns_class.av_uI/I` (cif\_core.dic 2.3)

Measure  $[\sum |u(\text{net } I)| / \sum |\text{net } I|]$  for all measured intensities in a reflection class.

The permitted range is  $[0.0, \infty)$ .

Related item: `_diffrn_reflns_class.av_sgI/I` (alternate).

[diffrn\_reflns\_class]

\* `_diffrn_reflns_class.code` (code)  
`_diffrn_reflns_class.code` (cif\_core.dic 2.3)

The code identifying a certain reflection class.

Examples: '1', 'm1', 's2'. [diffrn\_reflns\_class]

`_diffrn_reflns_class.d_res_high` (float)  
`_diffrn_reflns_class.d_res_high` (cif\_core.dic 2.3)

The smallest value in ångströms for the interplanar spacings for the reflections in each measured reflection class. This is called the highest resolution for this reflection class.

The permitted range is  $[0.0, \infty)$ . [diffrn\_reflns\_class]

`_diffrn_reflns_class.d_res_low` (float)  
`_diffrn_reflns_class.d_res_low` (cif\_core.dic 2.3)

The largest value in ångströms of the interplanar spacings for the reflections for each measured reflection class. This is called the lowest resolution for this reflection class.

The permitted range is  $[0.0, \infty)$ . [diffrn\_reflns\_class]

`_diffrn_reflns_class.description` (text)  
`_diffrn_reflns_class.description` (cif\_core.dic 2.3)

Description of each reflection class.

Examples: 'm=1 first order satellites',  
'HOL0 common projection reflections'. [diffrn\_reflns\_class]

`_diffrn_reflns_class.number` (int)  
`_diffrn_reflns_class.number` (cif\_core.dic 2.3)

The total number of measured intensities for each reflection class, excluding the systematic absences arising from centring translations.

The permitted range is  $[0, \infty)$ . [diffrn\_reflns\_class]

**DIFFRN\_SCALE\_GROUP**

Data items in the DIFFRN\_SCALE\_GROUP category record details of the scaling factors applied to place all intensities in the reflection lists on a common scale. Scaling groups might, for example, correspond to each film in a multi-film data set or each crystal in a multi-crystal data set.

Category group(s): `inclusive_group`  
`diffrn_group`

Category key(s): `_diffrn_scale_group.code`

*Example 1 – based on CAD-4 diffractometer data obtained for  $Yb(S-C_3H_4N)_2-(THF)_4$ .*

```
_diffrn_scale_group.code      A24
_diffrn_scale_group.I_net    1.021
```

\* `_diffrn_scale_group.code` (code)  
`_diffrn_scale_group.code` (cif\_core.dic 2.0.1)

The value of `_diffrn_scale_group.code` must uniquely identify a record in the DIFFRN\_SCALE\_GROUP list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

`_diffrn_reflns.scale_group_code`.

Examples: '1', '2', 'c1', 'c2'. [diffrn\_scale\_group]

`_diffrn_scale_group.I_net` (float)  
`_diffrn_scale_group.I_net` (cif\_core.dic 2.0.1)

The scale for a specific measurement group which is to be multiplied with the net intensity to place all intensities in the DIFFRN\_REFLN or REFLN list on a common scale.

The permitted range is  $[0.0, \infty)$ . [diffrn\_scale\_group]

**DIFFRN\_SOURCE**

Data items in the DIFFRN\_SOURCE category record details of the source of radiation used in the diffraction experiment.

Category group(s): `inclusive_group`  
`diffrn_group`

Category key(s): `_diffrn_source.diffrn_id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
_diffrn_source.diffrn_id    's1'
_diffrn_source.source      'rotating anode'
_diffrn_source.type        'Rigaku RU-200'
_diffrn_source.power       50
_diffrn_source.current     180
_diffrn_source.size        '8mm x 0.4 mm broad-focus'
```

`_diffrn_source.current` (float)  
`_diffrn_source.current` (cif\_core.dic 2.0.1)

The current in milliamperes at which the radiation source was operated.

[diffrn\_source]

**diffrn\_source.details** (text)

diffrn\_source\_details (cif\_core.dic 2.0.1)

A description of special aspects of the radiation source used.

[diffrn\_source]

\* **diffrn\_source.diffrn\_id**

This data item is a pointer to diffrn\_id in the DIFFRN category.

**diffrn\_source.power** (float)

diffrn\_source\_power (cif\_core.dic 2.0.1)

The power in kilowatts at which the radiation source was operated.

[diffrn\_source]

**diffrn\_source.size** (text)

diffrn\_source\_size (cif\_core.dic 2.0.1)

The dimensions of the source as viewed from the sample.

Examples: '8mm x 0.4 mm fine-focus', 'broad focus'. [diffrn\_source]

**diffrn\_source.source** (text)

diffrn\_radiation\_source (cifdic.c91 1.0)

diffrn\_source (cif\_core.dic 2.0)

The general class of the radiation source.

Examples: 'sealed X-ray tube', 'nuclear reactor', 'spallation source', 'electron microscope', 'rotating-anode X-ray tube', 'synchrotron'.

[diffrn\_source]

**diffrn\_source.take-off\_angle** (float)

diffrn\_source\_take-off\_angle (cif\_core.dic 2.3)

The complement of the angle in degrees between the normal to the surface of the X-ray tube target and the primary X-ray beam for beams generated by traditional X-ray tubes.

The permitted range is [0.00, 90.0].

Example: '1.5'.

[diffrn\_source]

**diffrn\_source.target** (code)

diffrn\_source\_target (cif\_core.dic 2.0.1)

The chemical element symbol for the X-ray target (usually the anode) used to generate X-rays. This can also be used for spallation sources.

The data value must be one of the following:

H He Li Be B C N O F Ne Na Mg Al Si P  
S Cl Ar K Ca Sc Ti V Cr Mn Fe Co Ni Cu Zn  
Ga Ge As Se Br Kr Rb Sr Y Zr Nb Mo Tc Ru Rh  
Pd Ag Cd In Sn Sb Te I Xe Cs Ba La Ce Pr Nd  
Pm Sm Eu Gd Tb Dy Ho Er Tm Yb Lu Hf Ta W Re  
Os Ir Pt Au Hg Tl Pb Bi Po At Rn Fr Ra Ac Th  
Pa U Np Pu Am Cm Bk Cf Es Fm Md No Lr

[diffrn\_source]

**diffrn\_source.type** (text)

diffrn\_source\_type (cif\_core.dic 2.0.1)

The make, model or name of the source of radiation.

Examples: 'NSLS beamline X8C', 'Rigaku RU200'.

[diffrn\_source]

**diffrn\_source.voltage** (float)

diffrn\_source\_voltage (cif\_core.dic 2.0.1)

The voltage in kilovolts at which the radiation source was operated.

[diffrn\_source]

**DIFFRN\_STANDARD\_REFLN**

Data items in the DIFFRN\_STANDARD\_REFLN category record details about the reflections treated as standards during the measurement of a set of diffraction intensities. Note that these are the individual standard reflections, not the results of the analysis of the standard reflections.

Category group(s): inclusive\_group  
diffrn\_group

Category key(s): diffrn\_standard\_refl.diffrn\_id  
diffrn\_standard\_refl.code

Example 2 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].

```
loop_
  diffrn_standard_refl.diffrn_id
  diffrn_standard_refl.code
  diffrn_standard_refl.index_h
  diffrn_standard_refl.index_k
  diffrn_standard_refl.index_l
s1 1 3 2 4
s1 1 1 9 1
s1 1 3 0 10
```

\* **diffrn\_standard\_refl.code** (code)

diffrn\_standard\_refl\_code (cif\_core.dic 2.0.1)

The code identifying a reflection measured as a standard reflection with the indices diffrn\_standard\_refl.index\_h, diffrn\_standard\_refl.index\_k and diffrn\_standard\_refl.index\_l. This is the same code as the diffrn\_refl.standard\_code in the DIFFRN\_REFLN list.

The following item(s) have an equivalent role in their respective categories:

diffrn\_refl.standard\_code.

Examples: '1', '2', 'c1', 'c2'.

[diffrn\_standard\_refl]

\* **diffrn\_standard\_refl.diffrn\_id**

This data item is a pointer to diffrn\_id in the DIFFRN category.

\* **diffrn\_standard\_refl.index\_h** (int)

diffrn\_standard\_refl\_index\_h (cif\_core.dic 2.0.1)

Miller index *h* of a standard reflection used in the diffraction measurement process.

[diffrn\_standard\_refl]

\* **diffrn\_standard\_refl.index\_k** (int)

diffrn\_standard\_refl\_index\_k (cif\_core.dic 2.0.1)

Miller index *k* of a standard reflection used in the diffraction measurement process.

[diffrn\_standard\_refl]

\* **diffrn\_standard\_refl.index\_l** (int)

diffrn\_standard\_refl\_index\_l (cif\_core.dic 2.0.1)

Miller index *l* of a standard reflection used in the diffraction measurement process.

[diffrn\_standard\_refl]

## DIFFRN\_STANDARDS

Data items in the DIFFRN\_STANDARDS category record details about the set of standard reflections used to monitor intensity stability during the measurement of diffraction intensities. Note that these records describe properties common to the set of standard reflections, not the standard reflections themselves.

Category group(s): **inclusive\_group**  
**diffrn\_group**

Category key(s): **\_diffrn\_standards.diffrn\_id**

*Example 1 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].*

```
_diffrn_standards.diffrn_id      's1'
_diffrn_standards.number         3
_diffrn_standards.interval_time  120
_diffrn_standards.decay_%        0
```

**\_diffrn\_standards.decay\_%** (float)  
**\_diffrn\_standards\_decay\_%** (cif\_core.dic 2.0.1)

The percentage decrease in the mean of the intensities for the set of standard reflections from the start of the measurement process to the end. This value usually affords a measure of the overall decay in crystal quality during the diffraction measurement process. Negative values are used in exceptional instances where the final intensities are greater than the initial ones.

The permitted range is  $(-\infty, 100.0]$ .

[diffrn\_standards]

\* **\_diffrn\_standards.diffrn\_id**

This data item is a pointer to **\_diffrn\_id** in the DIFFRN category.

**\_diffrn\_standards.interval\_count** (int)  
**\_diffrn\_standards\_interval\_count** (cif\_core.dic 2.0.1)

The number of reflection intensities between the measurement of standard reflection intensities.

The permitted range is  $[0, \infty)$ .

[diffrn\_standards]

**\_diffrn\_standards.interval\_time** (float)  
**\_diffrn\_standards\_interval\_time** (cif\_core.dic 2.0.1)

The time in minutes between the measurement of standard reflection intensities.

The permitted range is  $[0, \infty)$ .

[diffrn\_standards]

**\_diffrn\_standards.number** (int)  
**\_diffrn\_standards\_number** (cif\_core.dic 2.0.1)

The number of unique standard reflections used during the measurement of the diffraction intensities.

The permitted range is  $[0, \infty)$ .

[diffrn\_standards]

**\_diffrn\_standards.scale\_sigma** (float)  
**\_diffrn\_standards\_scale\_sigma** (cif\_core.dic 2.0.1)

The standard uncertainty (estimated standard deviation) of the individual mean standard scales applied to the intensity data.

The permitted range is  $[0.0, \infty)$ .

[diffrn\_standards]

**\_diffrn\_standards.scale\_u** (float)  
**\_diffrn\_standards\_scale\_u** (cif\_core.dic 2.3)

The standard uncertainty of the individual mean standard scales applied to the intensity data.

The permitted range is  $[0.0, \infty)$ .

Related item: **\_diffrn\_standards.scale\_sigma** (alternate).

[diffrn\_standards]

## ENTITY

Data items in the ENTITY category record details (such as chemical composition, name and source) about the molecular entities that are present in the crystallographic structure. Items in the various ENTITY subcategories provide a full chemical description of these molecular entities. Entities are of three types: polymer, non-polymer and water. Note that the water category includes only water; ordered solvent such as sulfate ion or acetone would be described as individual non-polymer entities. The ENTITY category is specific to macromolecular CIF applications and replaces the function of the CHEMICAL category in the CIF core. It is important to remember that the ENTITY data are not the result of the crystallographic experiment; those results are represented by the ATOM\_SITE data items. ENTITY data items describe the chemistry of the molecules under investigation and can most usefully be thought of as the ideal groups to which the structure is restrained or constrained during refinement. It is also important to remember that entities do not correspond directly to the enumeration of the contents of the asymmetric unit. Entities are described only once, even in those structures that contain multiple observations of an entity. The STRUCT\_ASYNC data items, which reference the entity list, describe and label the contents of the asymmetric unit.

Category group(s): **inclusive\_group**  
**entity\_group**

Category key(s): **\_entity.id**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
_entity.id
_entity.type
_entity.formula_weight
_entity.details
  1 polymer      10916
; The enzymatically competent form of HIV protease is a
  dimer. This entity corresponds to one monomer of an
  active dimer.
;
  2 non-polymer  647.2 .
  3 water        18 .
```

**\_entity.details** (text)

A description of special aspects of the entity.

[entity]

**\_entity.formula\_weight** (float)

Formula mass in daltons of the entity.

The permitted range is  $[1.0, \infty)$ .

[entity]

\* **\_entity.id** (code)

The value of **\_entity.id** must uniquely identify a record in the ENTITY list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

```
_atom_site.label_entity_id,
_entity_keywords.entity_id,
_entity_link.entity_id_1,
_entity_link.entity_id_2,
_entity_name_com.entity_id,
_entity_name_sys.entity_id,
_entity_poly.entity_id,
_entity_poly_seq.entity_id,
_entity_src_gen.entity_id,
_entity_src_nat.entity_id,
_struct_asym.entity_id,
_struct_ref.entity_id.
```

[entity]

**\_\_entity.src\_method** (ucode)

The method by which the sample for the entity was produced. Entities isolated directly from natural sources (tissues, soil samples *etc.*) are expected to have further information in the ENTITY\_SRC\_NAT category. Entities isolated from genetically manipulated sources are expected to have further information in the ENTITY\_SRC\_GEN category.

The data value must be one of the following:

```
nat      entity isolated from a natural source
man      entity isolated from a genetically manipulated source
syn      entity obtained synthetically
```

[entity]

**\_\_entity.type** (ucode)

Defines the type of the entity. Polymer entities are expected to have corresponding ENTITY\_POLY and associated entries. Non-polymer entities are expected to have corresponding CHEM\_COMP and associated entries. Water entities are not expected to have corresponding entries in the ENTITY category.

The data value must be one of the following:

```
polymer      entity is a polymer
non-polymer  entity is not a polymer
water        water in the solvent model
```

[entity]

**ENTITY\_KEYWORDS**

Data items in the ENTITY\_KEYWORDS category specify keywords relevant to the molecular entities. Note that this list of keywords is separate from the list that is used for the STRUCT\_BIOL data items and is intended to provide only the information that one would know about the molecular entity *if one did not know its structure*. Hence polypeptides are simply polypeptides, not cytokines or  $\beta$ - $\alpha$ -barrels, and polyribonucleic acids are simply poly-RNA, not transfer-RNA.

Category group(s): **inclusive\_group**  
**entity\_group**

Category key(s): **\_\_entity\_keywords.entity\_id**  
**\_\_entity\_keywords.text**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
__entity_keywords.entity_id
__entity_keywords.text
1 'polypeptide'
2 'natural product, inhibitor, reduced peptide'
```

**\*\_\_entity\_keywords.entity\_id**

This data item is a pointer to **\_\_entity.id** in the ENTITY category.

**\_\_entity\_keywords.text** (text)

Keywords describing this entity.

Examples: 'polypeptide', 'natural product', 'polysaccharide'.

[entity\_keywords]

**ENTITY\_LINK**

Data items in the ENTITY\_LINK category give details about the links between entities.

Category group(s): **inclusive\_group**  
**chem\_link\_group**

Category key(s): **\_\_entity\_link.link\_id**

**\_\_entity\_link.details** (text)

A description of special aspects of a link between chemical components in the structure.

[entity\_link]

**\*\_\_entity\_link.entity\_id\_1**

The entity ID of the first of the two entities joined by the link. This data item is a pointer to **\_\_entity.id** in the ENTITY category.

**\*\_\_entity\_link.entity\_id\_2**

The entity ID of the second of the two entities joined by the link. This data item is a pointer to **\_\_entity.id** in the ENTITY category.

**\_\_entity\_link.entity\_seq\_num\_1**

For a polymer entity, the sequence number in the first of the two entities containing the link. This data item is a pointer to **\_\_entity\_poly\_seq.num** in the ENTITY\_POLY\_SEQ category.

**\_\_entity\_link.entity\_seq\_num\_2**

For a polymer entity, the sequence number in the second of the two entities containing the link. This data item is a pointer to **\_\_entity\_poly\_seq.num** in the ENTITY\_POLY\_SEQ category.

**\*\_\_entity\_link.link\_id**

This data item is a pointer to **\_\_chem\_link.id** in the CHEM\_LINK category.

**ENTITY\_NAME\_COM**

Data items in the ENTITY\_NAME\_COM category record the common name or names associated with the entity. In some cases, the entity name may not be the same as the name of the biological structure. For example, haemoglobin  $\alpha$  chain would be the entity common name, not haemoglobin.

Category group(s): **inclusive\_group**  
**entity\_group**

Category key(s): **\_\_entity\_name\_com.entity\_id**  
**\_\_entity\_name\_com.name**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
__entity_name_com.entity_id
__entity_name_com.name
1 'HIV-1 protease monomer'
1 'HIV-1 PR monomer'
2 'acetyl-pepstatin'
2 'acetyl-Ile-Val-Asp-Statine-Ala-Ile-Statine'
3 'water'
```

**\*\_\_entity\_name\_com.entity\_id**

This data item is a pointer to **\_\_entity.id** in the ENTITY category.

**\*\_\_entity\_name\_com.name** (text)

A common name for the entity.

Examples: 'HIV protease monomer', 'hemoglobin alpha chain', '2-fluoro-1,4-dichloro benzene', 'arbutin'.

[entity\_name\_com]

**ENTITY\_NAME\_SYS**

Data items in the ENTITY\_NAME\_SYS category record the systematic name or names associated with the entity and the system that was used to construct the systematic name. In some cases, the entity name may not be the same as the name of the biological structure.

Category group(s): **inclusive\_group**

**entity\_group**

Category key(s): **\_entity\_name\_sys.entity\_id**

**\_entity\_name\_sys.name**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

loop\_

**\_entity\_name\_sys.entity\_id**

**\_entity\_name\_sys.name**

```
1 'EC 3.4.23.16'
2 'acetyl-Ile-Val-Asp-Sta-Ala-Ile-Sta'
3 water
```

\* **\_entity\_name\_sys.entity\_id**

This data item is a pointer to **\_entity.id** in the ENTITY category.

\* **\_entity\_name\_sys.name**

(text)

The systematic name for the entity.

Examples: 'hydroquinone-beta-D-pyranoside', 'EC 2.1.1.1',

'2-fluoro-1,4-dichlorobenzene'.

[entity\_name\_sys]

**\_entity\_name\_sys.system**

(text)

The system used to generate the systematic name of the entity.

Examples: 'Chemical Abstracts conventions', 'enzyme convention',

'Sigma catalog'.

[entity\_name\_sys]

**ENTITY\_POLY**

Data items in the ENTITY\_POLY category record details about the polymer, such as the type of the polymer, the number of monomers and whether it has nonstandard features.

Category group(s): **inclusive\_group**

**entity\_group**

Category key(s): **\_entity\_poly.entity\_id**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

loop\_

**\_entity\_poly.entity\_id**

**\_entity\_poly.type**

**\_entity\_poly.nstd\_chirality**

**\_entity\_poly.nstd\_linkage**

**\_entity\_poly.nstd\_monomer**

**\_entity\_poly.type\_details**

```
1 polypeptide(L) no no no .
```

\* **\_entity\_poly.entity\_id**

This data item is a pointer to **\_entity.id** in the ENTITY category.

**\_entity\_poly.nstd\_chirality**

(ucode)

A flag to indicate whether the polymer contains at least one monomer unit with chirality different from that specified in **\_entity\_poly.type**.

The data value must be one of the following:

no polymer contains no monomers with different chirality

n abbreviation for 'no'

yes polymer contains at least one monomer with different chirality

y abbreviation for 'yes'

[entity\_poly]

**\_entity\_poly.nstd\_linkage**

(ucode)

A flag to indicate whether the polymer contains at least one monomer-to-monomer link different from that implied by

**\_entity\_poly.type**.

The data value must be one of the following:

no polymer contains no different links

n abbreviation for 'no'

yes polymer contains at least one different link

y abbreviation for 'yes'

[entity\_poly]

**\_entity\_poly.nstd\_monomer**

(ucode)

A flag to indicate whether the polymer contains at least one monomer that is not considered standard.

The data value must be one of the following:

no polymer contains no nonstandard monomers

n abbreviation for 'no'

yes polymer contains at least one nonstandard monomer

y abbreviation for 'yes'

[entity\_poly]

**\_entity\_poly.number\_of\_monomers**

(int)

The number of monomers in the polymer.

The permitted range is [1, ∞).

[entity\_poly]

**\_entity\_poly.type**

(ucode)

The type of the polymer.

The data value must be one of the following:

polypeptide (D)

polypeptide (L)

polydeoxyribonucleotide

polyribonucleotide

polysaccharide (D)

polysaccharide (L)

other

[entity\_poly]

**\_entity\_poly.type\_details**

(text)

A description of special aspects of the polymer type.

Examples: 'monomer Ala 16 is a D-amino acid',

'the oligomer contains alternating RNA and DNA units'.

[entity\_poly]

**ENTITY\_POLY\_SEQ**

Data items in the ENTITY\_POLY\_SEQ category specify the sequence of monomers in a polymer. Allowance is made for the possibility of microheterogeneity in a sample by allowing a given sequence number to be correlated with more than one monomer ID. The corresponding ATOM\_SITE entries should reflect this heterogeneity.

Category group(s): **inclusive\_group**

**entity\_group**

Category key(s): **\_entity\_poly\_seq.entity\_id**

**\_entity\_poly\_seq.num**

**\_entity\_poly\_seq.mon\_id**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

loop\_

**\_entity\_poly\_seq.entity\_id**

**\_entity\_poly\_seq.num**

**\_entity\_poly\_seq.mon\_id**

```
1 1 PRO 1 2 GLN 1 3 ILE 1 4 THR 1 5 LEU
1 6 TRP 1 7 GLN 1 8 ARG 1 9 PRO 1 10 LEU
1 11 VAL 1 12 THR 1 13 ILE 1 14 LYS 1 15 ILE
1 16 GLY 1 17 GLY 1 18 GLN 1 19 LEU 1 20 LYS
1 21 GLU 1 22 ALA 1 23 LEU 1 24 LEU 1 25 ASP
```

# - - - data truncated for brevity - - -

**\*\_entity\_poly\_seq.entity\_id**

This data item is a pointer to `_entity.id` in the ENTITY category.

**\_entity\_poly\_seq.hetero** *(ucode)*

A flag to indicate whether this monomer in the polymer is heterogeneous in sequence. This would be rare.

The data value must be one of the following:

no	sequence is not heterogeneous at this monomer
n	abbreviation for 'no'
yes	sequence is heterogeneous at this monomer
y	abbreviation for 'yes'

Where no value is given, the assumed value is 'no'. [entity\_poly\_seq]

**\*\_entity\_poly\_seq.mon\_id**

This data item is a pointer to `_chem_comp.id` in the CHEM\_COMP category.

**\*\_entity\_poly\_seq.num** *(int)*

The value of `_entity_poly_seq.num` must uniquely and sequentially identify a record in the ENTITY\_POLY\_SEQ list. Note that this item must be a number and that the sequence numbers must progress in increasing numerical order.

The following item(s) have an equivalent role in their respective categories:

```

_atom_site.label_seq_id,
_entity_link.entity_seq_num 1,
_entity_link.entity_seq_num 2,
_geom_angle.atom_site_label_seq_id 1,
_geom_angle.atom_site_label_seq_id 2,
_geom_angle.atom_site_label_seq_id 3,
_geom_bond.atom_site_label_seq_id 1,
_geom_bond.atom_site_label_seq_id 2,
_geom_contact.atom_site_label_seq_id 1,
_geom_contact.atom_site_label_seq_id 2,
_geom_hbond.atom_site_label_seq_id A,
_geom_hbond.atom_site_label_seq_id D,
_geom_hbond.atom_site_label_seq_id H,
_geom_torsion.atom_site_label_seq_id 1,
_geom_torsion.atom_site_label_seq_id 2,
_geom_torsion.atom_site_label_seq_id 3,
_geom_torsion.atom_site_label_seq_id 4,
_struct_conf.beg_label_seq_id,
_struct_conf.end_label_seq_id,
_struct_conn.ptnr1_label_seq_id,
_struct_conn.ptnr2_label_seq_id,
_struct_mon_nucl.label_seq_id,
_struct_mon_prot.label_seq_id,
_struct_mon_prot_cis.label_seq_id,
_struct_ncs_dom_lim.beg_label_seq_id,
_struct_ncs_dom_lim.end_label_seq_id,
_struct_ref_seq.seq_align_beg,
_struct_ref_seq.seq_align_end,
_struct_ref_seq.dif.seq_num,
_struct_sheet_hbond.range_1_beg_label_seq_id,
_struct_sheet_hbond.range_1_end_label_seq_id,
_struct_sheet_hbond.range_2_beg_label_seq_id,
_struct_sheet_hbond.range_2_end_label_seq_id,
_struct_sheet_range.beg_label_seq_id,
_struct_sheet_range.end_label_seq_id,
_struct_site_gen.label_seq_id.

```

The permitted range is [1, ∞).

[entity\_poly\_seq]

**ENTITY\_SRC\_GEN**

Data items in the ENTITY\_SRC\_GEN category record details of the source from which the entity was obtained in cases where the source was genetically manipulated. The following are treated separately: items pertaining to the tissue from which the gene was obtained, items pertaining to the host organism for gene expression and items pertaining to the actual producing organism (plasmid).

Category group(s): **inclusive\_group**

**entity\_group**

Category key(s): **\_entity\_src\_gen.entity\_id**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```

loop_
_entity_src_gen.entity_id
_entity_src_gen.gene_src_common_name
_entity_src_gen.gene_src_genus
_entity_src_gen.gene_src_species
_entity_src_gen.gene_src_strain
_entity_src_gen.host_org_common_name
_entity_src_gen.host_org_genus
_entity_src_gen.host_org_species
_entity_src_gen.plasmid_name
1 'HIV-1' ? ? 'NY-5'
'bacteria' 'Escherichia' 'coli' 'pB322'

```

**\*\_entity\_src\_gen.entity\_id**

This data item is a pointer to `_entity.id` in the ENTITY category.

**\_entity\_src\_gen.gene\_src\_common\_name** *(text)*

The common name of the natural organism from which the gene was obtained.

Examples: 'man', 'yeast', 'bacteria'. [entity\_src\_gen]

**\_entity\_src\_gen.gene\_src\_details** *(text)*

A description of special aspects of the natural organism from which the gene was obtained.

[entity\_src\_gen]

**\_entity\_src\_gen.gene\_src\_genus** *(text)*

The genus of the natural organism from which the gene was obtained.

Examples: 'Homo', 'Saccharomyces', 'Escherichia'. [entity\_src\_gen]

**\_entity\_src\_gen.gene\_src\_species** *(text)*

The species of the natural organism from which the gene was obtained.

Examples: 'sapiens', 'cerevisiae', 'coli'. [entity\_src\_gen]

**\_entity\_src\_gen.gene\_src\_strain** *(text)*

The strain of the natural organism from which the gene was obtained, if relevant.

Examples: 'DH5a', 'BMH 71-18'. [entity\_src\_gen]

**\_entity\_src\_gen.gene\_src\_tissue** *(text)*

The tissue of the natural organism from which the gene was obtained.

Examples: 'heart', 'liver', 'eye lens'. [entity\_src\_gen]

**\_entity\_src\_gen.gene\_src\_tissue\_fraction** *(text)*

The subcellular fraction of the tissue of the natural organism from which the gene was obtained.

Examples: 'mitochondria', 'nucleus', 'membrane'. [entity\_src\_gen]

**\_entity\_src\_gen.host\_org\_common\_name** *(text)*

The common name of the organism that served as host for the production of the entity.

Examples: 'yeast', 'bacteria'. [entity\_src\_gen]

**\_\_entity\_src\_gen.host\_org\_details** (text)  
A description of special aspects of the organism that served as host for the production of the entity.

[entity\_src\_gen]

**\_\_entity\_src\_gen.host\_org\_genus** (text)  
The genus of the organism that served as host for the production of the entity.

Examples: 'Saccharomyces', 'Escherichia'.

[entity\_src\_gen]

**\_\_entity\_src\_gen.host\_org\_species** (text)  
The species of the organism that served as host for the production of the entity.

Examples: 'cerevisiae', 'coli'.

[entity\_src\_gen]

**\_\_entity\_src\_gen.host\_org\_strain** (text)  
The strain of the organism that served as host for the production of the entity.

Examples: 'DH5a', 'BMH 71-18'.

[entity\_src\_gen]

**\_\_entity\_src\_gen.plasmid\_details** (text)  
A description of special aspects of the plasmid that produced the entity in the host organism.

[entity\_src\_gen]

**\_\_entity\_src\_gen.plasmid\_name** (text)  
The name of the plasmid that produced the entity in the host organism.

Examples: 'pET3C', 'pT123sab'.

[entity\_src\_gen]

### ENTITY\_SRC\_NAT

Data items in the ENTITY\_SRC\_NAT category record details of the source from which the entity was obtained in cases where the entity was isolated directly from a natural tissue.

Category group(s): inclusive\_group  
entry\_group

Category key(s): \_\_entity\_src\_nat.entity\_id

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _entity_src_nat.entity_id
  _entity_src_nat.common_name
  _entity_src_nat.genus
  _entity_src_nat.species
  _entity_src_nat.details
  2 'bacteria' 'Actinomycetes' ?
  ; Acetyl-pepstatin was isolated by Dr. K. Oda, Osaka
    Prefecture University, and provided to us by Dr. Ben
    Dunn, University of Florida, and Dr. J. Kay, University
    of Wales.
  ;
```

\* **\_\_entity\_src\_nat.common\_name** (text)  
The common name of the organism from which the entity was isolated.

Examples: 'man', 'yeast', 'bacteria'.

[entity\_src\_nat]

**\_\_entity\_src\_nat.details** (text)  
A description of special aspects of the organism from which the entity was isolated.

[entity\_src\_nat]

\* **\_\_entity\_src\_nat.entity\_id** (text)  
This data item is a pointer to `__entity.id` in the ENTITY category.

\* **\_\_entity\_src\_nat.genus** (text)  
The genus of the organism from which the entity was isolated.

Examples: 'Homo', 'Saccharomyces', 'Escherichia'.

[entity\_src\_nat]

\* **\_\_entity\_src\_nat.species** (text)  
The species of the organism from which the entity was isolated.

Examples: 'sapiens', 'cerevisiae', 'coli'.

[entity\_src\_nat]

\* **\_\_entity\_src\_nat.strain** (text)  
The strain of the organism from which the entity was isolated.

Examples: 'DH5a', 'BMH 71-18'.

[entity\_src\_nat]

\* **\_\_entity\_src\_nat.tissue** (text)  
The tissue of the organism from which the entity was isolated.

Examples: 'heart', 'liver', 'eye lens'.

[entity\_src\_nat]

\* **\_\_entity\_src\_nat.tissue\_fraction** (text)  
The subcellular fraction of the tissue of the organism from which the entity was isolated.

Examples: 'mitochondria', 'nucleus', 'membrane'.

[entity\_src\_nat]

### ENTRY

There is only one item in the ENTRY category, `__entry.id`. This data item gives a name to this entry and is indirectly a key to the categories (such as CELL, GEOM, EXPTL) that describe information pertinent to the entire data block.

Category group(s): inclusive\_group

entry\_group

Category key(s): \_\_entry.id

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

<code>__entry.id</code>	'5HVP'
-------------------------	--------

*Example 2 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].*

<code>__entry.id</code>	'TOZ'
-------------------------	-------

\* **\_\_entry.id** (code)

`__audit_block_code` (cif\_core.dic 2.0.1)

The value of `__entry.id` identifies the data block. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

```
__atom_sites.entry_id,
__cell.entry_id,
__cell_measurement.entry_id,
__chemical.entry_id,
__chemical_formula.entry_id,
__computing.entry_id,
__database.entry_id,
__database_PDB_matrix.entry_id,
__entry_link.entry_id,
__exptl.entry_id,
__geom.entry_id,
__journal.entry_id,
__phasing_averaging.entry_id,
__phasing_isomorphous.entry_id,
__phasing_MAD.entry_id,
__phasing_MIR.entry_id,
__publ.entry_id,
__publ_manuscript_incl.entry_id,
```

`_refine.entry_id,`  
`_refine_analyze.entry_id,`  
`_reflns.entry_id,`  
`_struct.entry_id,`  
`_struct_keywords.entry_id,`  
`_struct_mon_details.entry_id,`  
`_symmetry.entry_id.` [entry]

**ENTRY\_LINK**

Data items in the ENTRY\_LINK category record the relationships between the current data block identified by `_entry.id` and other data blocks within the current file which may be referenced in the current data block.

Category group(s): `inclusive_group`  
`entry_group`

Category key(s): `_entry_link.id`  
`_entry_link.entry_id`

*Example 1 – example file for the one-dimensional incommensurately modulated structure of K<sub>2</sub>SeO<sub>4</sub>.*

```
loop_
  _entry_link.id
  _entry_link.entry_id
  _entry_link.details
  KSE_COM KSE_TEXT
    'experimental data common to ref./mod. structures'
  KSE_REF KSE_TEXT 'reference structure'
  KSE_MOD KSE_TEXT 'modulated structure'
```

`_entry_link.details` (text)

`_audit_link_block_description` (cif.core.dic 2.0.1)

A description of the relationship between the data blocks identified by `_entry_link.id` and `_entry_link.entry_id`.

[entry\_link]

\* `_entry_link.entry_id`

This data item is a pointer to `_entry.id` in the ENTRY category.

\* `_entry_link.id` (code)

`_audit_link_block_code` (cif.core.dic 2.0.1)

The value of `_entry_link.id` identifies a data block related to the current data block.

[entry\_link]

**EXPTL**

Data items in the EXPTL category record details about the experimental work prior to the intensity measurements and details about the absorption-correction technique employed.

Category group(s): `inclusive_group`  
`exptl_group`

Category key(s): `_exptl.entry_id`

*Example 1 – based on laboratory records for Yb(S-C<sub>5</sub>H<sub>4</sub>N)<sub>2</sub>(THF)<sub>4</sub>.*

```
_exptl.entry_id          datablock1
_exptl. absorpt_coefficient_mu  1.22
_exptl. absorpt_correction_T_max  0.896
_exptl. absorpt_correction_T_min  0.802
_exptl. absorpt_correction_type  integration
_exptl. absorpt_process_details
; Gaussian grid method from SHELX76
  Sheldrick, G. M., "SHELX-76: structure determination and
  refinement program", Cambridge University, UK, 1976
;
_exptl. crystals_number    1
_exptl. details
; Enraf-Nonius LT2 liquid nitrogen variable-temperature
  device used
;
_exptl. method             'single-crystal x-ray diffraction'
_exptl. method_details
; graphite monochromatized Cu K(alpha) fixed tube and
  Enraf-Nonius CAD4 diffractometer used
;
```

`_exptl. absorpt_coefficient_mu` (float)  
`_exptl. absorpt_coefficient_mu` (cif.core.dic 2.0.1)

The absorption coefficient  $\mu$  in reciprocal millimetres calculated from the atomic content of the cell, the density and the radiation wavelength.

The permitted range is [0.0,  $\infty$ ). [exptl]

`_exptl. absorpt_correction_T_max` (float)  
`_exptl. absorpt_correction_T_max` (cif.core.dic 2.0.1)

The maximum transmission factor for the crystal and radiation. The maximum and minimum transmission factors are also referred to as the absorption correction  $A$  or  $1/A^*$ .

The permitted range is [0.0, 1.0]. [exptl]

`_exptl. absorpt_correction_T_min` (float)  
`_exptl. absorpt_correction_T_min` (cif.core.dic 2.0.1)

The minimum transmission factor for the crystal and radiation. The maximum and minimum transmission factors are also referred to as the absorption correction  $A$  or  $1/A^*$ .

The permitted range is [0.0, 1.0]. [exptl]

`_exptl. absorpt_correction_type` (ucode)  
`_exptl. absorpt_correction_type` (cif.core.dic 2.0.1)

The absorption correction type and method. The value 'empirical' should *not* be used unless more detailed information is not available.

The data value must be one of the following:

analytical	analytical from crystal shape
cylinder	cylindrical
empirical	empirical from intensities
gaussian	Gaussian from crystal shape
integration	integration from crystal shape
multi-scan	symmetry-related measurements
none	no correction applied
numerical	numerical from crystal shape
psi-scan	$\psi$ -scan corrections
refdelf	refined from $\Delta F$
sphere	spherical

[exptl]

`_exptl. absorpt_process_details` (text)  
`_exptl. absorpt_process_details` (cif.core.dic 2.0.1)

Description of the absorption process applied to the intensities. A literature reference should be supplied for  $\psi$ -scan techniques.

Example: 'Tompas analytical'. [exptl]

`_exptl. crystals_number` (int)  
`_exptl. crystals_number` (cif.core.dic 2.0.1)

The total number of crystals used in the measurement of intensities.

The permitted range is [1,  $\infty$ ). [exptl]

`_exptl. details` (text)  
`_exptl. special_details` (cif.core.dic 2.0.1)

Any special information about the experimental work prior to the intensity measurement. See also `_exptl_crystal_preparation`.

[exptl]

\* `_exptl.entry_id`

This data item is a pointer to `_entry.id` in the ENTRY category.

\* `_exptl.method` (line)

The method used in the experiment.

Examples: 'single-crystal x-ray diffraction',  
'single-crystal neutron diffraction',  
'single-crystal electron diffraction', 'fiber x-ray diffraction',  
'fiber neutron diffraction', 'fiber electron diffraction',  
'single-crystal joint x-ray and neutron diffraction',  
'single-crystal joint x-ray and electron diffraction',  
'solution nmr', 'solid-state nmr', 'theoretical model', 'other'.

[exptl]

**exptl.method\_details**

(text)

A description of special aspects of the experimental method.

Examples: '29 structures', 'minimized average structure'. [exptl]

**EXPTL\_CRYSTAL**

Data items in the EXPTL\_CRYSTAL category record the results of experimental measurements on the crystal or crystals used, such as shape, size or density.

Category group(s): inclusive\_group

exptl\_group

Category key(s): \_exptl\_crystal.id

*Example 1 – based on laboratory records for Yb(S-C<sub>5</sub>H<sub>4</sub>N)<sub>2</sub>(THF)<sub>4</sub>.*

```

_exptl_crystal.id          xst21
_exptl_crystal.colour      'pale yellow'
_exptl_crystal.density_diffn 1.113
_exptl_crystal.density_Matthews 1.01
_exptl_crystal.density_meas 1.11
_exptl_crystal.density_meas_temp 294.5
_exptl_crystal.density_method 'neutral buoyancy'
_exptl_crystal.density_percent_sol 0.15
                                # P = 1 - (1.23*N*MMass) / V
_exptl_crystal.description 'hexagonal rod, uncut'
_exptl_crystal.F_000        202
_exptl_crystal.preparation
; hanging drop, crystal soaked in 10% ethylene glycol for
; 10 h, then placed in nylon loop at data collection time
;
_exptl_crystal.size_max     0.30
_exptl_crystal.size_mid    0.20
_exptl_crystal.size_min    0.05
_exptl_crystal.size_rad    0.025

```

*Example 2 - using separate items to define upper and lower limits for a value.*

```

_exptl_crystal.density_meas_gt 2.5
_exptl_crystal.density_meas_lt 5.0

```

*Example 3 - here the density was measured at some unspecified temperature below room temperature.*

```

_exptl_crystal.density_meas_temp_lt 300

```

**exptl\_crystal.colour**

(line)

*exptl\_crystal.colour* (cif.core.dic 2.0.1)

The colour of the crystal.

Example: 'dark green'. [exptl\_crystal]

**exptl\_crystal.colour\_lustre**

(line)

*exptl\_crystal.colour\_lustre* (cif.core.dic 2.3)The enumeration list of standardized names developed for the International Centre for Diffraction Data. The colour of a crystal is given by the combination of *exptl\_crystal.colour\_modifier* with *exptl\_crystal.colour\_primary*, as in 'dark-green' or 'bluish-violet', if necessary combined with *exptl\_crystal.colour\_lustre*, as in 'metallic-green'.Related item: *exptl\_crystal.colour* (alternate).

The data value must be one of the following:

```

metallic
dull
clear

```

[exptl\_crystal]

**exptl\_crystal.colour\_modifier**

(line)

*exptl\_crystal.colour\_modifier* (cif.core.dic 2.3)The enumeration list of standardized names developed for the International Centre for Diffraction Data. The colour of a crystal is given by the combination of *exptl\_crystal.colour\_modifier* with *exptl\_crystal.colour\_primary*, as in 'dark-green' or 'bluish-violet', if necessary combined with *exptl\_crystal.colour\_lustre*, as in 'metallic-green'.Related item: *exptl\_crystal.colour* (alternate).

The data value must be one of the following:

```

light
dark
whitish
blackish
grayish
brownish
reddish
pinkish
orangish
yellowish
greenish
bluish

```

[exptl\_crystal]

**exptl\_crystal.colour\_primary**

(line)

*exptl\_crystal.colour\_primary* (cif.core.dic 2.3)The enumeration list of standardized names developed for the International Centre for Diffraction Data. The colour of a crystal is given by the combination of *exptl\_crystal.colour\_modifier* with *exptl\_crystal.colour\_primary*, as in 'dark-green' or 'bluish-violet', if necessary combined with *exptl\_crystal.colour\_lustre*, as in 'metallic-green'.Related item: *exptl\_crystal.colour* (alternate).

The data value must be one of the following:

```

colourless
white
black
gray
brown
red
pink
orange
yellow
green
blue
violet

```

[exptl\_crystal]

**exptl\_crystal.density\_diffn**

(float)

*exptl\_crystal.density\_diffn* (cif.core.dic 2.0.1)

Density values calculated from the crystal cell and contents. The units are megagrams per cubic metre (grams per cubic centimetre).

The permitted range is [0.0, ∞).

[exptl\_crystal]

**exptl\_crystal.density\_Matthews**

(float)

The density of the crystal, expressed as the ratio of the volume of the asymmetric unit to the molecular mass of a monomer of the structure, in units of ångströms<sup>3</sup> per dalton.Reference: Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.

[exptl\_crystal]

**exptl\_crystal.density\_meas**

(float, su)

*exptl\_crystal.density\_meas* (cif.core.dic 2.3)

Density values measured using standard chemical and physical methods. The units are megagrams per cubic metre (grams per cubic centimetre).

The permitted range is [0.0, ∞).

Related item: *exptl\_crystal.density\_meas\_esd* (associated esd).

[exptl\_crystal]

**exptl\_crystal.density\_meas\_esd**

(float)

The standard uncertainty (estimated standard deviation) of *exptl\_crystal.density\_meas*.Related item: *exptl\_crystal.density\_meas* (associated value).

[exptl\_crystal]

**\_exptl\_crystal.density\_meas\_gt** (float)  
*\_exptl\_crystal.density\_meas\_gt* (cif\_core.dic 2.3)

The value above which the density measured using standard chemical and physical methods lies. The units are megagrams per cubic metre (grams per cubic centimetre). *\_exptl\_crystal.density\_meas\_gt* and *\_exptl\_crystal.density\_meas\_lt* should not be used to report new experimental work, for which *\_exptl\_crystal.density\_meas* should be used. These items are intended for use in reporting information in existing databases and archives which would be misleading if reported under *\_exptl\_crystal.density\_meas*.

The permitted range is [0.0, ∞).

Related item: *\_exptl\_crystal.density\_meas* (alternate).

Example: '2.5' (lower limit for the density (only the range within which the density lies was given in the original paper)). [exptl\_crystal]

**\_exptl\_crystal.density\_meas\_lt** (float)  
*\_exptl\_crystal.density\_meas\_lt* (cif\_core.dic 2.3)

The value below which the density measured using standard chemical and physical methods lies. The units are megagrams per cubic metre (grams per cubic centimetre). *\_exptl\_crystal.density\_meas\_gt* and *\_exptl\_crystal.density\_meas\_lt* should not be used to report new experimental work, for which *\_exptl\_crystal.density\_meas* should be used. These items are intended for use in reporting information in existing databases and archives which would be misleading if reported under *\_exptl\_crystal.density\_meas*.

The permitted range is [0.0, ∞).

Related item: *\_exptl\_crystal.density\_meas* (alternate).

Examples: '1.0' (specimen floats in water), '5.0' (upper limit for the density (only the range within which the density lies was given in the original paper)). [exptl\_crystal]

**\_exptl\_crystal.density\_meas\_temp** (float, su)  
*\_exptl\_crystal.density\_meas\_temp* (cif\_core.dic 2.3)

Temperature in kelvins at which *\_exptl\_crystal.density\_meas* was determined.

The permitted range is [0.0, ∞).

[exptl\_crystal]

**\_exptl\_crystal.density\_meas\_temp\_esd** (float)

The standard uncertainty (estimated standard deviation) of *\_exptl\_crystal.density\_meas\_temp*.

[exptl\_crystal]

**\_exptl\_crystal.density\_meas\_temp\_gt** (float)  
*\_exptl\_crystal.density\_meas\_temp\_gt* (cif\_core.dic 2.3)

Temperature in kelvins above which *\_exptl\_crystal.density\_meas* was determined. *\_exptl\_crystal.density\_meas\_temp\_gt* and *\_exptl\_crystal.density\_meas\_temp\_lt* should not be used for reporting new work, for which the correct temperature of measurement should be given. These items are intended for use in reporting information stored in databases or archives which would be misleading if reported under *\_exptl\_crystal.density\_meas\_temp*.

The permitted range is [0.0, ∞).

Related item: *\_exptl\_crystal.density\_meas\_temp* (alternate).

[exptl\_crystal]

**\_exptl\_crystal.density\_meas\_temp\_lt** (float)  
*\_exptl\_crystal.density\_meas\_temp\_lt* (cif\_core.dic 2.3)

Temperature in kelvins below which *\_exptl\_crystal.density\_meas* was determined. *\_exptl\_crystal.density\_meas\_temp\_gt* and *\_exptl\_crystal.density\_meas\_temp\_lt* should not be used for reporting new work, for which the correct temperature of measurement should be given. These items are intended for use in reporting information stored in databases

or archives which would be misleading if reported under *\_exptl\_crystal.density\_meas\_temp*.

The permitted range is [0.0, ∞).

Related item: *\_exptl\_crystal.density\_meas\_temp* (alternate).

Example: '300' (The density was measured at some unspecified temperature below room temperature.). [exptl\_crystal]

**\_exptl\_crystal.density\_method** (text)  
*\_exptl\_crystal.density\_method* (cif\_core.dic 2.0.1)

The method used to measure *\_exptl\_crystal.density\_meas*.

[exptl\_crystal]

**\_exptl\_crystal.density\_percent\_sol** (float)

Density value  $P$  calculated from the crystal cell and contents, expressed as per cent solvent.

$$P = 1 - (1.23NM_{\text{Mass}})/V,$$

where  $N$  = the number of molecules in the unit cell,  $M_{\text{Mass}}$  = the molecular mass of each molecule ( $\text{g mol}^{-1}$ ),  $V$  = the volume of the unit cell ( $\text{\AA}^3$ ) and 1.23 = a conversion factor evaluated as

$$\frac{(0.74\text{cm}^3/\text{g})(10^{24}\text{\AA}^3/\text{cm}^3)}{(6.02 \times 10^{23}\text{molecules/mole})},$$

where 0.74 is an assumed value for the partial specific volume of the molecule.

The permitted range is [0.0, ∞).

[exptl\_crystal]

**\_exptl\_crystal.description** (text)

*\_exptl\_crystal.description* (cif\_core.dic 2.0.1)

A description of the quality and habit of the crystal. The crystal dimensions should not normally be reported here; use instead the specific items in the EXPTL\_CRYSTAL category relating to size for the gross dimensions of the crystal and data items in the EXPTL\_CRYSTAL\_FACE category to describe the relationship between individual faces.

[exptl\_crystal]

**\_exptl\_crystal.F\_000** (int)

*\_exptl\_crystal.F\_000* (cif\_core.dic 2.0.1)

The effective number of electrons in the crystal unit cell contributing to  $F(000)$ . This may contain dispersion contributions and is calculated as

$$F(000) = \left[ \left( \sum f_r \right)^2 + \left( \sum f_i \right)^2 \right]^{1/2},$$

where  $f_r$  = real part of the scattering factors at  $\theta = 0^\circ$ ,  $f_i$  = imaginary part of the scattering factors at  $\theta = 0^\circ$  and the sum is taken over each atom in the unit cell.

The permitted range is [1, ∞).

[exptl\_crystal]

\* **\_exptl\_crystal.id** (code)

*\_exptl\_crystal.id* (cif\_core.dic 2.0.1)

The value of *\_exptl\_crystal.id* must uniquely identify a record in the EXPTL\_CRYSTAL list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

*\_diffrn.crystal\_id*,

*\_exptl\_crystal\_grow.crystal\_id*,

*\_exptl\_crystal\_face.crystal\_id*,

*\_exptl\_crystal\_grow\_comp.crystal\_id*,

*\_refln.crystal\_id*.

[exptl\_crystal]

**\_exptl\_crystal.preparation** (text)  
 \_exptl\_crystal\_preparation (cif\_core.dic 2.0.1)

Details of crystal growth and preparation of the crystal (e.g. mounting) prior to the intensity measurements.

Example: 'mounted in an argon-filled quartz capillary'.

[exptl\_crystal]

**\_exptl\_crystal.size\_max** (float)  
 \_exptl\_crystal\_size\_max (cif\_core.dic 2.0.1)

The maximum dimension of the crystal. This item may appear in a list with `_exptl_crystal.id` if multiple crystals are used in the experiment.

The permitted range is [0.0, ∞).

[exptl\_crystal]

**\_exptl\_crystal.size\_mid** (float)  
 \_exptl\_crystal\_size\_mid (cif\_core.dic 2.0.1)

The medial dimension of the crystal. This item may appear in a list with `_exptl_crystal.id` if multiple crystals are used in the experiment.

The permitted range is [0.0, ∞).

[exptl\_crystal]

**\_exptl\_crystal.size\_min** (float)  
 \_exptl\_crystal\_size\_min (cif\_core.dic 2.0.1)

The minimum dimension of the crystal. This item may appear in a list with `_exptl_crystal.id` if multiple crystals are used in the experiment.

The permitted range is [0.0, ∞).

[exptl\_crystal]

**\_exptl\_crystal.size\_rad** (float)  
 \_exptl\_crystal\_size\_rad (cif\_core.dic 2.0.1)

The radius of the crystal, if the crystal is a sphere or a cylinder. This item may appear in a list with `_exptl_crystal.id` if multiple crystals are used in the experiment.

The permitted range is [0.0, ∞).

[exptl\_crystal]

### EXPTL\_CRYSTAL\_FACE

Data items in the EXPTL\_CRYSTAL\_FACE category record details of the crystal faces.

Category group(s): inclusive\_group  
 exptl\_group

Category key(s): \_exptl\_crystal\_face.crystal\_id  
 \_exptl\_crystal\_face.index\_h  
 \_exptl\_crystal\_face.index\_k  
 \_exptl\_crystal\_face.index\_l

Example 1 – based on laboratory records for  $\text{Yb}(\text{S-C}_5\text{H}_4\text{N})_2(\text{THF})_4$  for the 100 face of crystal xst11.

_exptl_crystal_face.crystal_id	xst11
_exptl_crystal_face.index_h	1
_exptl_crystal_face.index_k	0
_exptl_crystal_face.index_l	0
_exptl_crystal_face.diffraction_chi	42.56
_exptl_crystal_face.diffraction_kappa	30.23
_exptl_crystal_face.diffraction_phi	-125.56
_exptl_crystal_face.diffraction_psi	-0.34
_exptl_crystal_face.perp_dist	0.025

\*\_exptl\_crystal\_face.crystal\_id

This data item is a pointer to `_exptl_crystal.id` in the EXPTL\_CRYSTAL category.

**\_exptl\_crystal\_face.diffraction\_chi** (float)  
 \_exptl\_crystal\_face\_diffraction\_chi (cif\_core.dic 2.0.1)

The  $\chi$  diffractometer setting angle in degrees for a specific crystal face associated with `_exptl_crystal_face.perp_dist`.

[exptl\_crystal\_face]

**\_exptl\_crystal\_face.diffraction\_kappa** (float)  
 \_exptl\_crystal\_face\_diffraction\_kappa (cif\_core.dic 2.0.1)

The  $\kappa$  diffractometer setting angle in degrees for a specific crystal face associated with `_exptl_crystal_face.perp_dist`.

[exptl\_crystal\_face]

**\_exptl\_crystal\_face.diffraction\_phi** (float)  
 \_exptl\_crystal\_face\_diffraction\_phi (cif\_core.dic 2.0.1)

The  $\varphi$  diffractometer setting angle in degrees for a specific crystal face associated with `_exptl_crystal_face.perp_dist`.

[exptl\_crystal\_face]

**\_exptl\_crystal\_face.diffraction\_psi** (float)  
 \_exptl\_crystal\_face\_diffraction\_psi (cif\_core.dic 2.0.1)

The  $\psi$  diffractometer setting angle in degrees for a specific crystal face associated with `_exptl_crystal_face.perp_dist`.

[exptl\_crystal\_face]

\*\_exptl\_crystal\_face.index\_h (int)  
 \_exptl\_crystal\_face\_index\_h (cif\_core.dic 2.0.1)

Miller index  $h$  of the crystal face associated with the value `_exptl_crystal_face.perp_dist`.

[exptl\_crystal\_face]

\*\_exptl\_crystal\_face.index\_k (int)  
 \_exptl\_crystal\_face\_index\_k (cif\_core.dic 2.0.1)

Miller index  $k$  of the crystal face associated with the value `_exptl_crystal_face.perp_dist`.

[exptl\_crystal\_face]

\*\_exptl\_crystal\_face.index\_l (int)  
 \_exptl\_crystal\_face\_index\_l (cif\_core.dic 2.0.1)

Miller index  $l$  of the crystal face associated with the value `_exptl_crystal_face.perp_dist`.

[exptl\_crystal\_face]

**\_exptl\_crystal\_face.perp\_dist** (float)  
 \_exptl\_crystal\_face\_perp\_dist (cif\_core.dic 2.0.1)

The perpendicular distance in millimetres from the face to the centre of rotation of the crystal.

The permitted range is [0.0, ∞).

[exptl\_crystal\_face]

### EXPTL\_CRYSTAL\_GROW

Data items in the EXPTL\_CRYSTAL\_GROW category record details about the conditions and methods used to grow the crystal.

Category group(s): inclusive\_group  
 exptl\_group

Category key(s): \_exptl\_crystal\_grow.crystal\_id

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

_exptl_crystal_grow.crystal_id	1
_exptl_crystal_grow.method	'hanging drop'
_exptl_crystal_grow.apparatus	'Linbro plates'
_exptl_crystal_grow.atmosphere	'room air'
_exptl_crystal_grow.pH	4.7
_exptl_crystal_grow.temp	18 (3)
_exptl_crystal_grow.time	'approximately 2 days'

**\_exptl\_crystal\_grow.apparatus** (text)

The physical apparatus in which the crystal was grown.

Examples: 'Linbro plate', 'sandwich box', 'ACA plates'.

[exptl\_crystal\_grow]

**\_exptl\_crystal\_grow.atmosphere** (text)  
The nature of the gas or gas mixture in which the crystal was grown.  
Examples: 'room air', 'nitrogen', 'argon'. [exptl\_crystal\_grow]

\* **\_exptl\_crystal\_grow.crystal\_id**  
This data item is a pointer to \_exptl\_crystal.id in the EXPTL\_CRYSTAL category.

**\_exptl\_crystal\_grow.details** (text)  
A description of special aspects of the crystal growth.  
Examples:  
; Solution 2 was prepared as a well solution and mixed. A droplet containing 2 \ml of solution 1 was delivered onto a cover slip; 2 \ml of solution 2 was added to the droplet without mixing.  
;  
; Crystal plates were originally stored at room temperature for 1 week but no nucleation occurred. They were then transferred to 4 degrees C, at which temperature well formed single crystals grew in 2 days.  
;  
; The dependence on pH for successful crystal growth is very sharp. At pH 7.4 only showers of tiny crystals grew, at pH 7.5 well formed single crystals grew, at pH 7.6 no crystallization occurred at all.  
; [exptl\_crystal\_grow]

**\_exptl\_crystal\_grow.method** (text)  
The method used to grow the crystals.  
Examples: 'batch precipitation', 'batch dialysis', 'hanging drop vapor diffusion', 'sitting drop vapor diffusion'. [exptl\_crystal\_grow]

**\_exptl\_crystal\_grow.method\_ref** (text)  
A literature reference that describes the method used to grow the crystals.  
Example: 'McPherson et al., 1988'. [exptl\_crystal\_grow]

**\_exptl\_crystal\_grow.pH** (float)  
The pH at which the crystal was grown. If more than one pH was employed during the crystallization process, the final pH should be noted here and the protocol involving multiple pH values should be described in \_exptl\_crystal\_grow.details.  
The permitted range is [0.0, ∞).  
Examples: '7.4', '7.6', '4.3'. [exptl\_crystal\_grow]

**\_exptl\_crystal\_grow.pressure** (float, su)  
The ambient pressure in kilopascals at which the crystal was grown.  
The permitted range is [0.0, ∞).  
Related item: \_exptl\_crystal\_grow.pressure\_esd (associated esd). [exptl\_crystal\_grow]

**\_exptl\_crystal\_grow.pressure\_esd** (float)  
The standard uncertainty (estimated standard deviation) of \_exptl\_crystal\_grow.pressure.  
Related item: \_exptl\_crystal\_grow.pressure (associated value). [exptl\_crystal\_grow]

**\_exptl\_crystal\_grow.seeding** (text)  
A description of the protocol used for seeding the crystal growth.  
Examples: 'macroseeding',  
; Microcrystals were introduced from a previous crystal growth experiment by transfer with a human hair.  
; [exptl\_crystal\_grow]

**\_exptl\_crystal\_grow.seeding\_ref** (text)  
A literature reference that describes the protocol used to seed the crystal.  
Example: 'Stura et al., 1989'. [exptl\_crystal\_grow]

**\_exptl\_crystal\_grow.temp** (float, su)  
The temperature in kelvins at which the crystal was grown. If more than one temperature was employed during the crystallization process, the final temperature should be noted here and the protocol involving multiple temperatures should be described in \_exptl\_crystal\_grow.details.  
The permitted range is [0.0, ∞).  
Related item: \_exptl\_crystal\_grow.temp\_esd (associated esd). [exptl\_crystal\_grow]

**\_exptl\_crystal\_grow.temp\_details** (text)  
A description of special aspects of temperature control during crystal growth. [exptl\_crystal\_grow]

**\_exptl\_crystal\_grow.temp\_esd** (float)  
The standard uncertainty (estimated standard deviation) of \_exptl\_crystal\_grow.temp.  
Related item: \_exptl\_crystal\_grow.temp (associated value). [exptl\_crystal\_grow]

**\_exptl\_crystal\_grow.time** (text)  
The approximate time that the crystal took to grow to the size used for data collection.  
Examples: 'overnight', '2-4 days', '6 months'. [exptl\_crystal\_grow]

### EXPTL\_CRYSTAL\_GROW\_COMP

Data items in the EXPTL\_CRYSTAL\_GROW\_COMP category record details about the components of the solutions that were 'mixed' (by whatever means) to produce the crystal. In general, solution 1 is the solution that contains the molecule to be crystallized and solution 2 is the solution that contains the precipitant. However, the number of solutions required to describe the crystallization protocol is not limited to 2. Details of the crystallization protocol should be given in \_exptl\_crystal\_grow\_comp.details using the solutions described in EXPTL\_CRYSTAL\_GROW\_COMP.

Category group(s): **inclusive\_group**

**exptl\_group**

Category key(s): \_exptl\_crystal\_grow\_comp.id

\_exptl\_crystal\_grow\_comp.crystal\_id

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _exptl_crystal_grow_comp.crystal_id
  _exptl_crystal_grow_comp.id
  _exptl_crystal_grow_comp.sol_id
  _exptl_crystal_grow_comp.name
  _exptl_crystal_grow_comp.volume
  _exptl_crystal_grow_comp.conc
  _exptl_crystal_grow_comp.details
  1 1 1 'HIV-1 protease' '0.002 ml' '6 mg/ml'
  ; The protein solution was in a buffer containing 25 mM NaCl,
  100 mM NaMES/ MES buffer, pH 7.5, 3 mM NaAzide
  ;
  1 2 2 'NaCl' '0.200 ml' '4 M' 'in 3 mM NaAzide'
  1 3 2 'Acetic Acid' '0.047 ml' '100 mM' 'in 3 mM NaAzide'
  1 4 2 'Na Acetate' '0.053 ml' '100 mM'
  ; in 3 mM NaAzide. Buffer components were mixed to produce a
  pH of 4.7 according to a ratio calculated from the pKa. The
  actual pH of solution 2 was not measured.
  ;
  1 5 2 'water' '0.700 ml' 'neat' 'in 3 mM NaAzide'
```

**\_exptl\_crystal\_grow\_comp.conc** (line)  
The concentration of the solution component.  
Examples: '200 \ml', '0.1 ml'. [exptl\_crystal\_grow\_comp]

\***\_exptl\_crystal\_grow\_comp.crystal\_id**  
This data item is a pointer to **\_exptl\_crystal\_id** in the EXPTL\_CRYSTAL category.

**\_exptl\_crystal\_grow\_comp.details** (text)  
A description of any special aspects of the solution component. When the solution component is the one that contains the macromolecule, this could be the specification of the buffer in which the macromolecule was stored. When the solution component is a buffer component, this could be the methods (or formula) used to achieve a desired pH.  
Examples: 'in 3 mM NaAzide',  
; The protein solution was in a buffer  
containing 25 mM NaCl, 100 mM NaMES/MES  
buffer, pH 7.5, 3 mM NaAzide  
;  
; in 3 mM NaAzide. Buffer components were mixed  
to produce a pH of 4.7 according to a ratio  
calculated from the pKa. The actual pH of  
solution 2 was not measured.  
; [exptl\_crystal\_grow\_comp]

\***\_exptl\_crystal\_grow\_comp.id** (line)  
The value of **\_exptl\_crystal\_grow\_comp.id** must uniquely identify each item in the EXPTL\_CRYSTAL\_GROW\_COMP list. Note that this item need not be a number; it can be any unique identifier.  
Examples: '1', 'A', 'protein in buffer'. [exptl\_crystal\_grow\_comp]

**\_exptl\_crystal\_grow\_comp.name** (line)  
A common name for the component of the solution.  
Examples: 'protein in buffer', 'acetic acid'. [exptl\_crystal\_grow\_comp]

**\_exptl\_crystal\_grow\_comp.sol\_id** (line)  
An identifier for the solution to which the given solution component belongs.  
Examples: '1', 'well solution', 'solution A'. [exptl\_crystal\_grow\_comp]

**\_exptl\_crystal\_grow\_comp.volume** (line)  
The volume of the solution component.  
Examples: '200 \ml', '0.1 ml'. [exptl\_crystal\_grow\_comp]

## GEOM

Data items in the GEOM and related (GEOM\_ANGLE, GEOM\_BOND, GEOM\_CONTACT, GEOM\_HBOND and GEOM\_TORSION) categories record details about the molecular geometry as calculated from the contents of the ATOM, CELL and SYMMETRY data. Geometry data are therefore redundant, in that they can be calculated from other more fundamental quantities in the data block. However, they provide a check on the correctness of both sets of data and enable the most important geometric data to be identified for publication by setting the appropriate publication flag.

Category group(s): **inclusive\_group**  
**geom\_group**  
Category key(s): **\_geom.entry\_id**

**\_geom.details** (text)  
**\_geom\_special\_details** (cif\_core.dic 2.0.1)  
A description of geometry not covered by the existing data names in the GEOM categories, such as least-squares planes.  
[geom]

\***\_geom.entry\_id**  
This data item is a pointer to **\_entry.id** in the ENTRY category.

## GEOM\_ANGLE

Data items in the GEOM\_ANGLE category record details about the bond angles as calculated from the contents of the ATOM, CELL and SYMMETRY data.

Category group(s): **inclusive\_group**  
**geom\_group**  
Category key(s): **\_geom\_angle.atom\_site\_id\_1**  
**\_geom\_angle.atom\_site\_id\_2**  
**\_geom\_angle.atom\_site\_id\_3**  
**\_geom\_angle.site\_symmetry\_1**  
**\_geom\_angle.site\_symmetry\_2**  
**\_geom\_angle.site\_symmetry\_3**

Example 1 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].

```
loop_
  _geom_angle.atom_site_id_1
  _geom_angle.atom_site_id_2
  _geom_angle.atom_site_id_3
  _geom_angle.value
  _geom_angle.value_esd
  _geom_angle.site_symmetry_1
  _geom_angle.site_symmetry_2
  _geom_angle.site_symmetry_3
  _geom_angle.publ_flag
  C2 O1 C5 111.6 0.2 1_555 1_555 1_555 yes
  O1 C2 C3 110.9 0.2 1_555 1_555 1_555 yes
  O1 C2 O21 122.2 0.3 1_555 1_555 1_555 yes
  C3 C2 O21 127.0 0.3 1_555 1_555 1_555 yes
  C2 C3 N4 101.3 0.2 1_555 1_555 1_555 yes
  C2 C3 C31 111.3 0.2 1_555 1_555 1_555 yes
  C2 C3 H3 107 1 1_555 1_555 1_555 no
  N4 C3 C31 116.7 0.2 1_555 1_555 1_555 yes
# - - - data truncated for brevity - - -
```

**\_geom\_angle.atom\_site\_auth\_asym\_id\_1**  
An optional identifier of the first of the three atom sites that define the angle. This data item is a pointer to **\_atom\_site.auth\_asym\_id** in the ATOM\_SITE category.

**\_geom\_angle.atom\_site\_auth\_asym\_id\_2**  
An optional identifier of the second of the three atom sites that define the angle. This data item is a pointer to **\_atom\_site.auth\_asym\_id** in the ATOM\_SITE category.

**\_geom\_angle.atom\_site\_auth\_asym\_id\_3**  
An optional identifier of the third of the three atom sites that define the angle. This data item is a pointer to **\_atom\_site.auth\_asym\_id** in the ATOM\_SITE category.

**\_geom\_angle.atom\_site\_auth\_atom\_id\_1**  
An optional identifier of the first of the three atom sites that define the angle. This data item is a pointer to **\_atom\_site.auth\_atom\_id** in the ATOM\_SITE category.

**\_geom\_angle.atom\_site\_auth\_atom\_id\_2**  
An optional identifier of the second of the three atom sites that define the angle. This data item is a pointer to **\_atom\_site.auth\_atom\_id** in the ATOM\_SITE category.



geom\_angle.atom\_site\_label\_seq\_id\_3

An optional identifier of the third of the three atom sites that define the angle. This data item is a pointer to atom\_site.label\_seq\_id in the ATOM\_SITE category.

geom\_angle.publ\_flag (ucode)  
geom\_angle.publ\_flag (cif-core.dic 2.0.1)

This code signals whether the angle is referred to in a publication or should be placed in a table of significant angles.

The data value must be one of the following:

- no do not include angle in special list
- n abbreviation for 'no'
- yes do include angle in special list
- y abbreviation for 'yes'

[geom\_angle]

\* geom\_angle.site\_symmetry\_1 (symop)  
geom\_angle.site\_symmetry\_1 (cif-core.dic 2.0.1)

The symmetry code of the first of the three atom sites that define the angle.

Where no value is given, the assumed value is '1\_555'.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),

'7\_645' (7th symmetry position: +a on x, -b on y). [geom\_angle]

\* geom\_angle.site\_symmetry\_2 (symop)  
geom\_angle.site\_symmetry\_2 (cif-core.dic 2.0.1)

The symmetry code of the second of the three atom sites that define the angle.

Where no value is given, the assumed value is '1\_555'.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),

'7\_645' (7th symmetry position: +a on x, -b on y). [geom\_angle]

\* geom\_angle.site\_symmetry\_3 (symop)  
geom\_angle.site\_symmetry\_3 (cif-core.dic 2.0.1)

The symmetry code of the third of the three atom sites that define the angle.

Where no value is given, the assumed value is '1\_555'.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),

'7\_645' (7th symmetry position: +a on x, -b on y). [geom\_angle]

geom\_angle.value (float, su)  
geom\_angle (cif-core.dic 2.0.1)

Angle in degrees defined by the three sites geom\_angle.atom\_site\_id\_1, geom\_angle.atom\_site\_id\_2 and geom\_angle.atom\_site\_id\_3.

Related item: geom\_angle.value\_esd (associated esd). [geom\_angle]

geom\_angle.value\_esd (float)  
The standard uncertainty (estimated standard deviation) of geom\_angle.value.

Related item: geom\_angle.value (associated value). [geom\_angle]

**GEOM\_BOND**

Data items in the GEOM\_BOND category record details about the bond lengths as calculated from the contents of the ATOM, CELL and SYMMETRY data.

Category group(s): inclusive\_group

geom\_group

Category key(s): geom\_bond.atom\_site\_id\_1  
geom\_bond.atom\_site\_id\_2  
geom\_bond.site\_symmetry\_1  
geom\_bond.site\_symmetry\_2

Example 1 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].

```
loop
  _geom_bond.atom_site_id_1
  _geom_bond.atom_site_id_2
  _geom_bond.dist
  _geom_bond.dist_esd
  _geom_bond.site_symmetry_1
  _geom_bond.site_symmetry_2
  _geom_bond.publ_flag
  O1 C2 1.342 0.004 1_555 1_555 yes
  O1 C5 1.439 0.003 1_555 1_555 yes
  C2 C3 1.512 0.004 1_555 1_555 yes
  C2 O21 1.199 0.004 1_555 1_555 yes
  C3 N4 1.465 0.003 1_555 1_555 yes
  C3 C31 1.537 0.004 1_555 1_555 yes
  C3 H3 1.00 0.03 1_555 1_555 no
  N4 C5 1.472 0.003 1_555 1_555 yes
# - - - data truncated for brevity - - -
```

geom\_bond.atom\_site\_auth\_asym\_id\_1

An optional identifier of the first of the two atom sites that define the bond. This data item is a pointer to atom\_site.auth\_asym\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_auth\_asym\_id\_2

An optional identifier of the second of the two atom sites that define the bond. This data item is a pointer to atom\_site.auth\_asym\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_auth\_atom\_id\_1

An optional identifier of the first of the two atom sites that define the bond. This data item is a pointer to atom\_site.auth\_atom\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_auth\_atom\_id\_2

An optional identifier of the second of the two atom sites that define the bond. This data item is a pointer to atom\_site.auth\_atom\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_auth\_comp\_id\_1

An optional identifier of the first of the two atom sites that define the bond. This data item is a pointer to atom\_site.auth\_comp\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_auth\_comp\_id\_2

An optional identifier of the second of the two atom sites that define the bond. This data item is a pointer to atom\_site.auth\_comp\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_auth\_seq\_id\_1

An optional identifier of the first of the two atom sites that define the bond. This data item is a pointer to atom\_site.auth\_seq\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_auth\_seq\_id\_2

An optional identifier of the second of the two atom sites that define the bond. This data item is a pointer to atom\_site.auth\_seq\_id in the ATOM\_SITE category.

\* geom\_bond.atom\_site\_id\_1

geom\_bond\_atom\_site\_label\_1 (cif\_core.dic 2.0.1)

The identifier of the first of the two atom sites that define the bond. This data item is a pointer to atom\_site.id in the ATOM\_SITE category.

\* geom\_bond.atom\_site\_id\_2

geom\_bond\_atom\_site\_label\_2 (cif\_core.dic 2.0.1)

The identifier of the second of the two atom sites that define the bond. This data item is a pointer to atom\_site.id in the ATOM\_SITE category.

geom\_bond.atom\_site\_label\_alt\_id\_1

An optional identifier of the first of the two atom sites that define the bond. This data item is a pointer to atom\_site.label\_alt\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_label\_alt\_id\_2

An optional identifier of the second of the two atom sites that define the bond. This data item is a pointer to atom\_site.label\_alt\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_label\_asym\_id\_1

An optional identifier of the first of the two atom sites that define the bond. This data item is a pointer to atom\_site.label\_asym\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_label\_asym\_id\_2

An optional identifier of the second of the two atom sites that define the bond. This data item is a pointer to atom\_site.label\_asym\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_label\_atom\_id\_1

An optional identifier of the first of the two atom sites that define the bond. This data item is a pointer to atom\_site.label\_atom\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_label\_atom\_id\_2

An optional identifier of the second of the two atom sites that define the bond. This data item is a pointer to atom\_site.label\_atom\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_label\_comp\_id\_1

An optional identifier of the first of the two atom sites that define the bond. This data item is a pointer to atom\_site.label\_comp\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_label\_comp\_id\_2

An optional identifier of the second of the two atom sites that define the bond. This data item is a pointer to atom\_site.label\_comp\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_label\_seq\_id\_1

An optional identifier of the first of the two atom sites that define the bond. This data item is a pointer to atom\_site.label\_seq\_id in the ATOM\_SITE category.

geom\_bond.atom\_site\_label\_seq\_id\_2

An optional identifier of the second of the two atom sites that define the bond. This data item is a pointer to atom\_site.label\_seq\_id in the ATOM\_SITE category.

geom\_bond.dist

(float, su)

geom\_bond\_distance (cif\_core.dic 2.0.1)

The intramolecular bond distance in ångströms.

The permitted range is [0.0, ∞).

Related item: geom\_bond.dist\_esd (associated esd).

[geom\_bond]

geom\_bond.dist\_esd

(float)

The standard uncertainty (estimated standard deviation) of geom\_bond.dist.

Related item: geom\_bond.dist (associated value).

[geom\_bond]

geom\_bond.publ\_flag

(ucode)

geom\_bond\_publ\_flag (cif\_core.dic 2.0.1)

This code signals whether the bond distance is referred to in a publication or should be placed in a list of significant bond distances.

The data value must be one of the following:

- no do not include bond in special list
- n abbreviation for 'no'
- yes do include bond in special list
- y abbreviation for 'yes'

[geom\_bond]

\* geom\_bond.site\_symmetry\_1

(symop)

geom\_bond\_site\_symmetry\_1 (cif\_core.dic 2.0.1)

The symmetry code of the first of the two atom sites that define the bond.

Where no value is given, the assumed value is '1\_555'.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),

'7\_645' (7th symmetry position: +a on x, -b on y).

[geom\_bond]

\* geom\_bond.site\_symmetry\_2

(symop)

geom\_bond\_site\_symmetry\_2 (cif\_core.dic 2.0.1)

The symmetry code of the second of the two atom sites that define the bond.

Where no value is given, the assumed value is '1\_555'.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),

'7\_645' (7th symmetry position: +a on x, -b on y).

[geom\_bond]

geom\_bond.valence

(int)

geom\_bond\_valence (cif\_core.dic 2.3)

The bond valence calculated from geom\_bond.dist.

[geom\_bond]

**GEOM\_CONTACT**

Data items in the GEOM\_CONTACT category record details about interatomic contacts as calculated from the contents of the ATOM, CELL and SYMMETRY data.

Category group(s): **inclusive\_group**

**geom\_group**

Category key(s): **geom\_contact.atom\_site\_id\_1**  
**geom\_contact.atom\_site\_id\_2**  
**geom\_contact.site\_symmetry\_1**  
**geom\_contact.site\_symmetry\_2**

*Example 1 – based on data set CLPHO6 of Ferguson, Ruhl, McKervey & Browne [Acta Cryst. (1992), C48, 2262–2264].*

```
loop_
  _geom_contact.atom_site_id_1
  _geom_contact.atom_site_id_2
  _geom_contact.dist
  _geom_contact.dist_esd
  _geom_contact.site_symmetry_1
  _geom_contact.site_symmetry_2
  _geom_contact.publ_flag
  O(1) O(2) 2.735 0.003 . . yes
  H(01) O(2) 1.82 . . . no
```

**geom\_contact.atom\_site\_auth\_asym\_id\_1**

An optional identifier of the first of the two atom sites that define the contact. This data item is a pointer to **atom\_site.auth\_asym\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_auth\_asym\_id\_2**

An optional identifier of the second of the two atom sites that define the contact. This data item is a pointer to **atom\_site.auth\_asym\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_auth\_atom\_id\_1**

An optional identifier of the first of the two atom sites that define the contact. This data item is a pointer to **atom\_site.auth\_atom\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_auth\_atom\_id\_2**

An optional identifier of the second of the two atom sites that define the contact. This data item is a pointer to **atom\_site.auth\_atom\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_auth\_comp\_id\_1**

An optional identifier of the first of the two atom sites that define the contact. This data item is a pointer to **atom\_site.auth\_comp\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_auth\_comp\_id\_2**

An optional identifier of the second of the two atom sites that define the contact. This data item is a pointer to **atom\_site.auth\_comp\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_auth\_seq\_id\_1**

An optional identifier of the first of the two atom sites that define the contact. This data item is a pointer to **atom\_site.auth\_seq\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_auth\_seq\_id\_2**

An optional identifier of the second of the two atom sites that define the contact. This data item is a pointer to **atom\_site.auth\_seq\_id** in the ATOM\_SITE category.

**\*geom\_contact.atom\_site\_id\_1**

**geom\_contact\_atom\_site\_label\_1** (cif.core.dic 2.0.1)

The identifier of the first of the two atom sites that define the contact. This data item is a pointer to **atom\_site.id** in the ATOM\_SITE category.

**\*geom\_contact.atom\_site\_id\_2**

**geom\_contact\_atom\_site\_label\_2** (cif.core.dic 2.0.1)

The identifier of the second of the two atom sites that define the contact. This data item is a pointer to **atom\_site.id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_label\_alt\_id\_1**

An optional identifier of the first of the two atom sites that define the contact. This data item is a pointer to **atom\_site.label\_alt\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_label\_alt\_id\_2**

An optional identifier of the second of the two atom sites that define the contact. This data item is a pointer to **atom\_site.label\_alt\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_label\_asym\_id\_1**

An optional identifier of the first of the two atom sites that define the contact. This data item is a pointer to **atom\_site.label\_asym\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_label\_asym\_id\_2**

An optional identifier of the second of the two atom sites that define the contact. This data item is a pointer to **atom\_site.label\_asym\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_label\_atom\_id\_1**

An optional identifier of the first of the two atom sites that define the contact. This data item is a pointer to **atom\_site.label\_atom\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_label\_atom\_id\_2**

An optional identifier of the second of the two atom sites that define the contact. This data item is a pointer to **atom\_site.label\_atom\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_label\_comp\_id\_1**

An optional identifier of the first of the two atom sites that define the contact. This data item is a pointer to **atom\_site.label\_comp\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_label\_comp\_id\_2**

An optional identifier of the second of the two atom sites that define the contact. This data item is a pointer to **atom\_site.label\_comp\_id** in the ATOM\_SITE category.

**geom\_contact.atom\_site\_label\_seq\_id\_1**

An optional identifier of the first of the two atom sites that define the contact. This data item is a pointer to atom\_site.label\_seq\_id in the ATOM\_SITE category.

**geom\_contact.atom\_site\_label\_seq\_id\_2**

An optional identifier of the second of the two atom sites that define the contact. This data item is a pointer to atom\_site.label\_seq\_id in the ATOM\_SITE category.

**geom\_contact.dist** (float, su)

geom\_contact\_distance (cif.core.dic 2.0.1)

The interatomic contact distance in ångströms.

The permitted range is [0.0, ∞).

Related item: geom\_contact.dist\_esd (associated esd). [geom\_contact]

**geom\_contact.dist\_esd** (float)

The standard uncertainty (estimated standard deviation) of geom\_contact.dist.

Related item: geom\_contact.dist (associated value). [geom\_contact]

**geom\_contact.publ\_flag** (ucode)

geom\_contact\_publ\_flag (cif.core.dic 2.0.1)

This code signals whether the contact distance is referred to in a publication or should be placed in a list of significant contact distances.

The data value must be one of the following:

no	do not include distance in special list
n	abbreviation for 'no'
yes	do include distance in special list
y	abbreviation for 'yes'

[geom\_contact]

**\*geom\_contact.site\_symmetry\_1** (symop)

geom\_contact\_site\_symmetry\_1 (cif.core.dic 2.0.1)

The symmetry code of the first of the two atom sites that define the contact.

Where no value is given, the assumed value is '1\_555'.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),

'7\_645' (7th symmetry position: +a on x, -b on y). [geom\_contact]

**\*geom\_contact.site\_symmetry\_2** (symop)

geom\_contact\_site\_symmetry\_2 (cif.core.dic 2.0.1)

The symmetry code of the second of the two atom sites that define the contact.

Where no value is given, the assumed value is '1\_555'.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),

'7\_645' (7th symmetry position: +a on x, -b on y). [geom\_contact]

Example 1 – based on  $C_{14}H_{13}ClN_2O.H_2O$ , reported by Palmer, Puddle & Lisgarten [Acta Cryst. (1993), C49, 1777–1779].

```
loop_
  _geom_hbond.atom_site_id_D
  _geom_hbond.atom_site_id_H
  _geom_hbond.atom_site_id_A
  _geom_hbond.dist_DH
  _geom_hbond.dist_HA
  _geom_hbond.dist_DA
  _geom_hbond.angle_DHA
  _geom_hbond.publ_flag
N6  HN6  OW  0.888  1.921  2.801  169.6  yes
OW  HO2  O7  0.917  1.923  2.793  153.5  yes
OW  HO1  N10 0.894  1.886  2.842  179.7  yes
```

**geom\_hbond.angle\_DHA** (float, su)

geom\_hbond\_angle\_DHA (cif.core.dic 2.0.1)

The angle in degrees defined by the donor-, hydrogen- and acceptor-atom sites in a hydrogen bond.

The permitted range is [0.0, ∞).

Related item: geom\_hbond.angle\_DHA\_esd (associated esd). [geom\_hbond]

**geom\_hbond.angle\_DHA\_esd** (float)

The standard uncertainty (estimated standard deviation) of geom\_hbond.angle\_DHA.

Related item: geom\_hbond.angle\_DHA (associated value). [geom\_hbond]

**geom\_hbond.atom\_site\_auth\_asym\_id\_A**

An optional identifier of the acceptor-atom site that defines the hydrogen bond. This data item is a pointer to atom\_site.auth\_asym\_id in the ATOM\_SITE category.

**geom\_hbond.atom\_site\_auth\_asym\_id\_D**

An optional identifier of the donor-atom site that defines the hydrogen bond. This data item is a pointer to atom\_site.auth\_asym\_id in the ATOM\_SITE category.

**geom\_hbond.atom\_site\_auth\_asym\_id\_H**

An optional identifier of the hydrogen-atom site that defines the hydrogen bond. This data item is a pointer to atom\_site.auth\_asym\_id in the ATOM\_SITE category.

**geom\_hbond.atom\_site\_auth\_atom\_id\_A**

An optional identifier of the acceptor-atom site that defines the hydrogen bond. This data item is a pointer to atom\_site.auth\_atom\_id in the ATOM\_SITE category.

**geom\_hbond.atom\_site\_auth\_atom\_id\_D**

An optional identifier of the donor-atom site that defines the hydrogen bond. This data item is a pointer to atom\_site.auth\_atom\_id in the ATOM\_SITE category.

**geom\_hbond.atom\_site\_auth\_atom\_id\_H**

An optional identifier of the hydrogen-atom site that defines the hydrogen bond. This data item is a pointer to atom\_site.auth\_atom\_id in the ATOM\_SITE category.

**GEOM\_HBOND**

Data items in the GEOM\_HBOND category record details about hydrogen bonds as calculated from the contents of the ATOM, CELL and SYMMETRY data.

Category group(s): inclusive\_group

geom\_group

Category key(s): geom\_hbond.atom\_site\_id\_A

geom\_hbond.atom\_site\_id\_D

geom\_hbond.atom\_site\_id\_H

geom\_hbond.site\_symmetry\_A

geom\_hbond.site\_symmetry\_D

geom\_hbond.site\_symmetry\_H



**\_\_geom\_hbond.atom\_site\_label\_seq\_id\_D**  
 An optional identifier of the donor-atom site that defines the hydrogen bond. This data item is a pointer to `__atom_site.label_seq_id` in the ATOM\_SITE category.

**\_\_geom\_hbond.atom\_site\_label\_seq\_id\_H**  
 An optional identifier of the hydrogen-atom site that defines the hydrogen bond. This data item is a pointer to `__atom_site.label_seq_id` in the ATOM\_SITE category.

**\_\_geom\_hbond.dist\_DA** (float, su)  
**\_\_geom\_hbond\_distance\_DA** (cif\_core.dic 2.0.1)  
 The distance in ångströms between the donor- and acceptor-atom sites in a hydrogen bond.  
 The permitted range is [0.0, ∞).  
 Related item: `__geom_hbond.dist_DA_esd` (associated esd). [geom\_hbond]

**\_\_geom\_hbond.dist\_DA\_esd** (float)  
 The standard uncertainty (estimated standard deviation) in ångströms of `__geom_hbond.dist_DA`.  
 Related item: `__geom_hbond.dist_DH` (associated value). [geom\_hbond]

**\_\_geom\_hbond.dist\_DH** (float, su)  
**\_\_geom\_hbond\_distance\_DH** (cif\_core.dic 2.0.1)  
 The distance in ångströms between the donor- and hydrogen-atom sites in a hydrogen bond.  
 The permitted range is [0.0, ∞).  
 Related item: `__geom_hbond.dist_DH_esd` (associated esd). [geom\_hbond]

**\_\_geom\_hbond.dist\_DH\_esd** (float)  
 The standard uncertainty (estimated standard deviation) in ångströms of `__geom_hbond.dist_DH`.  
 Related item: `__geom_hbond.dist_DH` (associated value). [geom\_hbond]

**\_\_geom\_hbond.dist\_HA** (float, su)  
**\_\_geom\_hbond\_distance\_HA** (cif\_core.dic 2.0.1)  
 The distance in ångströms between the hydrogen- and acceptor-atom sites in a hydrogen bond.  
 The permitted range is [0.0, ∞).  
 Related item: `__geom_hbond.dist_HA_esd` (associated esd). [geom\_hbond]

**\_\_geom\_hbond.dist\_HA\_esd** (float)  
 The standard uncertainty (estimated standard deviation) in ångströms of `__geom_hbond.dist_HA`.  
 Related item: `__geom_hbond.dist_HA` (associated value). [geom\_hbond]

**\_\_geom\_hbond.publ\_flag** (ucode)  
**\_\_geom\_hbond\_publ\_flag** (cif\_core.dic 2.0.1)  
 This code signals whether the hydrogen-bond information is referred to in a publication or should be placed in a table of significant hydrogen-bond geometry.  
 The data value must be one of the following:

no	do not include bond in special list
n	abbreviation for 'no'
yes	do include bond in special list
y	abbreviation for 'yes'

[geom\_hbond]

**\* \_\_geom\_hbond.site\_symmetry\_A** (symop)  
**\_\_geom\_hbond\_site\_symmetry\_A** (cif\_core.dic 2.0.1)  
 The symmetry code of the acceptor-atom site that defines the hydrogen bond.  
 Where no value is given, the assumed value is '1\_555'.  
 Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied), '7\_645' (7th symmetry position: +a on x, -b on y). [geom\_hbond]

**\* \_\_geom\_hbond.site\_symmetry\_D** (symop)  
**\_\_geom\_hbond\_site\_symmetry\_D** (cif\_core.dic 2.0.1)  
 The symmetry code of the donor-atom site that defines the hydrogen bond.  
 Where no value is given, the assumed value is '1\_555'.  
 Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied), '7\_645' (7th symmetry position: +a on x, -b on y). [geom\_hbond]

**\* \_\_geom\_hbond.site\_symmetry\_H** (symop)  
**\_\_geom\_hbond\_site\_symmetry\_H** (cif\_core.dic 2.0.1)  
 The symmetry code of the hydrogen-atom site that defines the hydrogen bond.  
 Where no value is given, the assumed value is '1\_555'.  
 Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied), '7\_645' (7th symmetry position: +a on x, -b on y). [geom\_hbond]

**GEOM.TORSION**

Data items in the GEOM\_TORSION category record details about torsion angles as calculated from the contents of the ATOM, CELL and SYMMETRY data. The vector direction `__geom_torsion.atom_site_id_2` to `__geom_torsion.atom_site_id_3` is the viewing direction, and the torsion angle is the angle of twist required to superimpose the projection of the vector between site 2 and site 1 onto the projection of the vector between site 3 and site 4. Clockwise torsions are positive, anti-clockwise torsions are negative.

Reference: Klyne, W. & Prelog, V. (1960). *Experientia*, **16**, 521–523.

Category group(s): `inclusive_group`  
                   `geom_group`

Category key(s): `__geom_torsion.atom_site_id_1`  
                   `__geom_torsion.atom_site_id_2`  
                   `__geom_torsion.atom_site_id_3`  
                   `__geom_torsion.atom_site_id_4`  
                   `__geom_torsion.site_symmetry_1`  
                   `__geom_torsion.site_symmetry_2`  
                   `__geom_torsion.site_symmetry_3`  
                   `__geom_torsion.site_symmetry_4`

---

*Example 1 – based on data set CLPHO6 of Ferguson, Ruhl, McKervey & Browne [Acta Cryst. (1992), C48, 2262–2264].*

```

loop_
  __geom_torsion.atom_site_id_1
  __geom_torsion.atom_site_id_2
  __geom_torsion.atom_site_id_3
  __geom_torsion.atom_site_id_4
  __geom_torsion.value
  __geom_torsion.site_symmetry_1
  __geom_torsion.site_symmetry_2
  __geom_torsion.site_symmetry_3
  __geom_torsion.site_symmetry_4
  __geom_torsion.publ_flag
C(9)  O(2)  C(7)  C(2)    71.8  .  .  .  .  yes
C(7)  O(2)  C(9)  C(10) -168.0 .  .  .  2_666 yes
C(10) O(3)  C(8)  C(6)   -167.7 .  .  .  .  yes
C(8)  O(3)  C(10) C(9)   -69.7 .  .  .  2_666 yes
O(1)  C(1)   C(2)  C(3)  -179.5 .  .  .  .  no
O(1)  C(1)  C(2)  C(7)   -0.6  .  .  .  .  no
    
```



**geom\_torsion.atom\_site\_label\_alt\_id\_3**  
 An optional identifier of the third of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_alt\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_alt\_id\_4**  
 An optional identifier of the fourth of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_alt\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_asym\_id\_1**  
 An optional identifier of the first of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_asym\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_asym\_id\_2**  
 An optional identifier of the second of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_asym\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_asym\_id\_3**  
 An optional identifier of the third of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_asym\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_asym\_id\_4**  
 An optional identifier of the fourth of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_asym\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_atom\_id\_1**  
 An optional identifier of the first of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_atom\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_atom\_id\_2**  
 An optional identifier of the second of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_atom\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_atom\_id\_3**  
 An optional identifier of the third of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_atom\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_atom\_id\_4**  
 An optional identifier of the fourth of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_atom\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_comp\_id\_1**  
 An optional identifier of the first of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_comp\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_comp\_id\_2**  
 An optional identifier of the second of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_comp\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_comp\_id\_3**  
 An optional identifier of the third of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_comp\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_comp\_id\_4**  
 An optional identifier of the fourth of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_comp\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_seq\_id\_1**  
 An optional identifier of the first of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_seq\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_seq\_id\_2**  
 An optional identifier of the second of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_seq\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_seq\_id\_3**  
 An optional identifier of the third of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_seq\_id in the ATOM\_SITE category.

**geom\_torsion.atom\_site\_label\_seq\_id\_4**  
 An optional identifier of the fourth of the four atom sites that define the torsion angle. This data item is a pointer to atom\_site.label\_seq\_id in the ATOM\_SITE category.

**geom\_torsion.publ\_flag** *(ucode)*  
geom\_torsion.publ\_flag (*cif-core.dic 2.0.1*)  
 This code signals whether the torsion angle is referred to in a publication or should be placed in a table of significant torsion angles. The data value must be one of the following:

no	do not include angle in special list
n	abbreviation for 'no'
yes	do include angle in special list
y	abbreviation for 'yes'

[geom\_torsion]

**\* geom\_torsion.site\_symmetry\_1** *(symop)*  
geom\_torsion.site\_symmetry\_1 (*cif-core.dic 2.0.1*)  
 The symmetry code of the first of the four atom sites that define the torsion angle.  
 Where no value is given, the assumed value is '1.555'.  
 Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),  
 '7\_645' (7th symmetry position: +a on x, -b on y). [geom\_torsion]

**\* geom\_torsion.site\_symmetry\_2** *(symop)*  
geom\_torsion.site\_symmetry\_2 (*cif-core.dic 2.0.1*)  
 The symmetry code of the second of the four atom sites that define the torsion angle.  
 Where no value is given, the assumed value is '1.555'.  
 Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),  
 '7\_645' (7th symmetry position: +a on x, -b on y). [geom\_torsion]

\* **\_geom\_torsion.site\_symmetry\_3** (symop)  
 \_geom\_torsion\_site\_symmetry\_3 (cif\_core.dic 2.0.1)

The symmetry code of the third of the four atom sites that define the torsion angle.

Where no value is given, the assumed value is '1\_555'.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),

'7\_645' (7th symmetry position: +a on x, -b on y). [geom\_torsion]

\* **\_geom\_torsion.site\_symmetry\_4** (symop)  
 \_geom\_torsion\_site\_symmetry\_4 (cif\_core.dic 2.0.1)

The symmetry code of the fourth of the four atom sites that define the torsion angle.

Where no value is given, the assumed value is '1\_555'.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),

'7\_645' (7th symmetry position: +a on x, -b on y). [geom\_torsion]

**\_geom\_torsion.value** (float, su)  
 \_geom\_torsion (cif\_core.dic 2.0.1)

The value of the torsion angle in degrees.

Related item: \_geom\_torsion.value\_esd (associated esd). [geom\_torsion]

**\_geom\_torsion.value\_esd** (float)  
 The standard uncertainty (estimated standard deviation) of \_geom\_torsion.value.

Related item: \_geom\_torsion.value (associated value). [geom\_torsion]

**JOURNAL**

Data items in the JOURNAL category record details about the book-keeping by the journal staff when processing a data block submitted for publication. The creator of a data block will not normally specify these data. The data names are not defined in the dictionary because they are for journal use only.

Category group(s): inclusive\_group  
 iucr\_group

Category key(s): \_journal.entry\_id

Example 1 – based on Acta Cryst. file for entry HL0007 [Willis, Beckwith & Tozer (1991). Acta Cryst. C47, 2276–2277].

_journal.entry_id	'TOZ'
_journal.date_recd_electronic	1991-04-15
_journal.date_from_coeditor	1991-04-18
_journal.date_accepted	1991-04-18
_journal.date_printers_first	1991-08-07
_journal.date_proofs_out	1991-08-07
_journal.coeditor_code	HL0007
_journal.techeditor_code	C910963
_journal.coden_ASTM	ACSCEE
_journal.name_full	'Acta Crystallographica Section C'
_journal.year	1991
_journal.volume	47
_journal.issue	NOV91
_journal.page_first	2276
_journal.page_last	2277

**\_journal.coden\_ASTM** (line)  
 \_journal\_codan\_ASTM (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.coden\_Cambridge** (line)  
 \_journal\_codan\_Cambridge (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.coeditor\_address** (text)  
 \_journal\_coeditor\_address (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.coeditor\_code** (line)  
 \_journal\_coeditor\_code (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.coeditor\_email** (line)  
 \_journal\_coeditor\_email (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.coeditor\_fax** (line)  
 \_journal\_coeditor\_fax (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.coeditor\_name** (line)  
 \_journal\_coeditor\_name (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.coeditor\_notes** (text)  
 \_journal\_coeditor\_notes (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.coeditor\_phone** (line)  
 \_journal\_coeditor\_phone (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.data\_validation\_number** (code)  
 \_journal\_data\_validation\_number (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.date\_accepted** (yyyy-mm-dd)  
 \_journal\_date\_accepted (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.date\_from\_coeditor** (yyyy-mm-dd)  
 \_journal\_date\_from\_coeditor (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.date\_printers\_final** (yyyy-mm-dd)  
 \_journal\_date\_printers\_final (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.date\_printers\_first** (yyyy-mm-dd)  
 \_journal\_date\_printers\_first (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.date\_proofs\_in** (yyyy-mm-dd)  
 \_journal\_date\_proofs\_in (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

**\_journal.date\_proofs\_out** (yyyy-mm-dd)  
 \_journal\_date\_proofs\_out (cif\_core.dic 2.0.1)

Journal data items are defined by the journal staff.

[journal]

<b>_journal.date_recd_copyright</b> <i>_journal_date_recd_copyright</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(yyyy-mm-dd)	<b>_journal.techeditor_code</b> <i>_journal_techeditor_code</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)
	[journal]		[journal]
<b>_journal.date_recd_electronic</b> <i>_journal_date_recd_electronic</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(yyyy-mm-dd)	<b>_journal.techeditor_email</b> <i>_journal_techeditor_email</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)
	[journal]		[journal]
<b>_journal.date_recd_hard_copy</b> <i>_journal_date_recd_hard_copy</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(yyyy-mm-dd)	<b>_journal.techeditor_fax</b> <i>_journal_techeditor_fax</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)
	[journal]		[journal]
<b>_journal.date_to_coeditor</b> <i>_journal_date_to_coeditor</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(yyyy-mm-dd)	<b>_journal.techeditor_name</b> <i>_journal_techeditor_name</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)
	[journal]		[journal]
<b>*_journal.entry_id</b> This data item is a pointer to <i>_entry.id</i> in the ENTRY category.		<b>_journal.techeditor_notes</b> <i>_journal_techeditor_notes</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(text)
			[journal]
<b>_journal.issue</b> <i>_journal_issue</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)	<b>_journal.techeditor_phone</b> <i>_journal_techeditor_phone</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)
	[journal]		[journal]
<b>_journal.language</b> <i>_journal_language</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)	<b>_journal.volume</b> <i>_journal_volume</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)
	[journal]		[journal]
<b>_journal.name_full</b> <i>_journal_name_full</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)	<b>_journal.year</b> <i>_journal_year</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)
	[journal]		[journal]
<b>_journal.page_first</b> <i>_journal_page_first</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)		
	[journal]		
<b>_journal.page_last</b> <i>_journal_page_last</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)		
	[journal]		
<b>_journal.paper_category</b> <i>_journal_paper_category</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)		
	[journal]		
<b>_journal.suppl_publ_number</b> <i>_journal_suppl_publ_number</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)		
	[journal]		
<b>_journal.suppl_publ_pages</b> <i>_journal_suppl_publ_pages</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(line)		
	[journal]		
<b>_journal.techeditor_address</b> <i>_journal_techeditor_address</i> (cif_core.dic 2.0.1) Journal data items are defined by the journal staff.	(text)		
	[journal]		

JOURNAL_INDEX	
Data items in the JOURNAL_INDEX category are used to list terms used to generate the journal indexes. The creator of a data block will not normally specify these data items.	
Category group(s):	<i>inclusive_group</i> <i>iucr_group</i>
Category key(s):	<i>_journal_index.type</i> <i>_journal_index.term</i>
<i>Example 1 – based on a paper by Zhu, Reynolds, Klein &amp; Trudell [Acta Cryst. (1994), C50, 2067–2069].</i>	
loop_	
<i>_journal_index.type</i>	
<i>_journal_index.term</i>	
<i>_journal_index.subterm</i>	
O	C16H19NO4 .
S	alkaloids (-) -norcocaine
S	(-) -norcocaine .
S	
;	[2R,3S-(2\ b,3\ b)] -methyl
;	3-(benzoyloxy) -8-azabicyclo[3.2.1]octane-2-carboxylate
;	
;	

<b>_journal_index.subterm</b> <i>_journal_index_subterm</i> (cif_core.dic 2.0.1) Journal index data items are defined by the journal staff.	(line)
	[journal_index]
<b>_journal_index.term</b> <i>_journal_index_term</i> (cif_core.dic 2.0.1) Journal index data items are defined by the journal staff.	(line)
	[journal_index]

**journal\_index.type** (line)  
journal\_index.type (cif\_core.dic 2.0.1)  
 Journal index data items are defined by the journal staff.  
 [journal\_index]

### PHASING

Data items in the PHASING category record details about the phasing of the structure, listing the various methods used in the phasing process. Details about the application of each method are listed in the appropriate subcategories.

Category group(s): **inclusive\_group**  
**phasing\_group**  
 Category key(s): **phasing.method**

*Example 1 – hypothetical example.*

```
loop_
  phasing.method
  'mir'
  'averaging'
```

\* **phasing.method** (ucode)

A listing of the method or methods used to phase this structure.

Examples: 'abinitio' (phasing by *ab initio* methods), 'averaging' (phase improvement by averaging over multiple images of the structure), 'dm' (phasing by direct methods), 'isas' (phasing by iterative single-wavelength anomalous scattering), 'isir' (phasing by iterative single-wavelength isomorphous replacement), 'isomorphous' (phasing beginning with phases calculated from an isomorphous structure), 'mad' (phasing by multiple-wavelength anomalous dispersion), 'mir' (phasing by multiple isomorphous replacement), 'miras' (phasing by multiple isomorphous replacement with anomalous scattering), 'mr' (phasing by molecular replacement), 'sir' (phasing by single isomorphous replacement), 'siras' (phasing by single isomorphous replacement with anomalous scattering). [phasing]

### PHASING\_AVERAGING

Data items in the PHASING\_AVERAGING category record details about the phasing of the structure where methods involving averaging of multiple observations of the molecule in the asymmetric unit are involved.

Category group(s): **inclusive\_group**  
**phasing\_group**  
 Category key(s): **phasing\_averaging.entry\_id**

*Example 1 – hypothetical example.*

```
phasing_averaging.entry_id  'EXAMHYPO'
phasing_averaging.method
; Iterative threefold averaging alternating with phase
  extensions by 0.5 reciprocal lattice units per cycle.
;
phasing_averaging.details
; The position of the threefold axis was redetermined every
  five cycles.
;
```

**phasing\_averaging.details** (text)

A description of special aspects of the averaging process.

[phasing\_averaging]

\* **phasing\_averaging.entry\_id**

This data item is a pointer to entry.id in the ENTRY category.

**phasing\_averaging.method** (text)

A description of the phase-averaging phasing method used to phase this structure. Note that this is not the computer program used, which is described in the SOFTWARE category, but rather the method itself. This data item should be used to describe significant methodological options used within the phase-averaging program.

[phasing\_averaging]

### PHASING\_ISOMORPHOUS

Data items in the PHASING\_ISOMORPHOUS category record details about the phasing of the structure where a model isomorphous to the structure being phased was used to generate the initial phases.

Category group(s): **inclusive\_group**  
**phasing\_group**  
 Category key(s): **phasing\_isomorphous.entry\_id**

*Example 1 – based on PDB entry 4PHV and laboratory records for the structure corresponding to PDB entry 4PHV.*

```
phasing_isomorphous.parent      'PDB entry 5HVP'
phasing_isomorphous.details
; The inhibitor and all solvent atoms were removed from the
  parent structure before beginning refinement. All static
  disorder present in the parent structure was also removed.
;
```

**phasing\_isomorphous.details** (text)

A description of special aspects of the isomorphous phasing.

Example:

```
; Residues 13-18 were eliminated from the starting model as it
  was anticipated that binding of the inhibitor would cause a
  structural rearrangement in this part of the structure.
;
```

[phasing\_isomorphous]

\* **phasing\_isomorphous.entry\_id**

This data item is a pointer to entry.id in the ENTRY category.

**phasing\_isomorphous.method** (text)

A description of the isomorphous-phasing method used to phase this structure. Note that this is not the computer program used, which is described in the SOFTWARE category, but rather the method itself. This data item should be used to describe significant methodological options used within the isomorphous phasing program.

Example:

```
; Iterative threefold averaging alternating with phase
  extension by 0.5 reciprocal lattice units per cycle.
;
```

[phasing\_isomorphous]

**phasing\_isomorphous.parent** (text)

Reference to the structure used to generate starting phases if the structure referenced in this data block was phased by virtue of being isomorphous to a known structure (e.g. a mutant that crystallizes in the same space group as the wild-type protein.)

[phasing\_isomorphous]

### PHASING\_MAD

Data items in the PHASING\_MAD category record details about the phasing of the structure where methods involving multiple-wavelength anomalous-dispersion techniques are involved.

Category group(s): **inclusive\_group**  
**phasing\_group**  
 Category key(s): **phasing\_MAD.entry\_id**

*Example 1 – based on a paper by Shapiro et al. [Nature (London) (1995), 374, 327–337].*

```
phasing_MAD.entry_id      'NCAD'
```

**phasing\_MAD.details** (text)

A description of special aspects of the MAD phasing.

[phasing\_MAD]

\* **phasing\_MAD.entry\_id**

This data item is a pointer to entry.id in the ENTRY category.

**phasing\_MAD.method** (text)  
 A description of the MAD phasing method used to phase this structure. Note that this is not the computer program used, which is described in the SOFTWARE category, but rather the method itself. This data item should be used to describe significant methodological options used within the MAD phasing program.  
 [phasing\_MAD]

**PHASING\_MAD\_CLUSTER**

Data items in the PHASING\_MAD\_CLUSTER category record details about a cluster of experiments that contributed to the generation of a set of phases.

Category group(s): inclusive\_group  
                   phasing\_group

Category key(s): phasing\_MAD\_clust.expt\_id  
                   phasing\_MAD\_clust.id

---

*Example 1 – based on a paper by Shapiro et al. [Nature (London) (1995), 374, 327–337].*

```
loop_
  _phasing_MAD_clust.id
  _phasing_MAD_clust.expt_id
  _phasing_MAD_clust.number_set
  '4 wavelength' 1 4
  '5 wavelength' 1 5
  '5 wavelength' 2 5
```

\* **phasing\_MAD\_clust.expt\_id**  
 This data item is a pointer to phasing\_MAD\_expt.id in the PHASING\_MAD\_EXPT category.

\* **phasing\_MAD\_clust.id** (code)  
 The value of phasing\_MAD\_clust.id must, together with phasing\_MAD\_clust.expt\_id, uniquely identify a record in the PHASING\_MAD\_CLUSTER list. Note that this item need not be a number; it can be any unique identifier.

*The following item(s) have an equivalent role in their respective categories:*  
phasing\_MAD\_set.clust\_id,  
phasing\_MAD\_ratio.clust\_id [phasing\_MAD\_clust]

**phasing\_MAD\_clust.number\_set** (int)  
 The number of data sets in this cluster of data sets.  
 [phasing\_MAD\_clust]

**PHASING\_MAD\_EXPT**

Data items in the PHASING\_MAD\_EXPT category record details about a MAD phasing experiment, such as the number of experiments that were clustered together to produce a set of phases or the statistics for those phases.

Category group(s): inclusive\_group  
                   phasing\_group

Category key(s): phasing\_MAD\_expt.id

---

*Example 1 – based on a paper by Shapiro et al. [Nature (London) (1995), 374, 327–337].*

```
loop_
  _phasing_MAD_expt.id
  _phasing_MAD_expt.number_clust
  _phasing_MAD_expt.R_normal_all
  _phasing_MAD_expt.R_normal_anom_sc
  _phasing_MAD_expt.delta_delta_phi
  _phasing_MAD_expt.delta_phi_sigma
  _phasing_MAD_expt.mean_fom
  1 2 0.063 0.451 58.5 20.3 0.88
  2 1 0.051 0.419 36.8 18.2 0.93
```

**phasing\_MAD\_expt.delta\_delta\_phi** (float)  
 The difference between two independent determinations of phasing\_MAD\_expt.delta\_phi.  
 [phasing\_MAD\_expt]

**phasing\_MAD\_expt.delta\_phi** (float)  
 The phase difference between  $F_t(h)$ , the structure factor due to normal scattering from all atoms, and  $F_a(h)$ , the structure factor due to normal scattering from only the anomalous scatterers.  
 Related item: phasing\_MAD\_expt.delta\_phi\_sigma (associated esd).  
 [phasing\_MAD\_expt]

**phasing\_MAD\_expt.delta\_phi\_sigma** (float)  
 The standard uncertainty (estimated standard deviation) of phasing\_MAD\_expt.delta\_phi.  
 Related item: phasing\_MAD\_expt.delta\_phi (associated value).  
 [phasing\_MAD\_expt]

\* **phasing\_MAD\_expt.id** (code)  
 The value of phasing\_MAD\_expt.id must uniquely identify each record in the PHASING\_MAD\_EXPT list.

*The following item(s) have an equivalent role in their respective categories:*  
phasing\_MAD\_clust.expt\_id,  
phasing\_MAD\_set.expt\_id,  
phasing\_MAD\_ratio.expt\_id [phasing\_MAD\_expt]

**phasing\_MAD\_expt.mean\_fom** (float)  
 The mean figure of merit.  
 [phasing\_MAD\_expt]

**phasing\_MAD\_expt.number\_clust** (int)  
 The number of clusters of data sets in this phasing experiment.  
 [phasing\_MAD\_expt]

**phasing\_MAD\_expt.R\_normal\_all** (float)  
 [phasing\_MAD\_expt]

**phasing\_MAD\_expt.R\_normal\_anom\_sc** (float)  
 [phasing\_MAD\_expt]

**PHASING\_MAD\_RATIO**

Data items in the PHASING\_MAD\_RATIO category record the ratios of phasing statistics between pairs of data sets in a MAD phasing experiment, in given shells of resolution.

Category group(s): inclusive\_group  
                   phasing\_group

Category key(s): phasing\_MAD\_ratio.clust\_id  
                   phasing\_MAD\_ratio.expt\_id  
                   phasing\_MAD\_ratio.wavelength\_1  
                   phasing\_MAD\_ratio.wavelength\_2

Example 1 – based on a paper by Shapiro et al. [Nature (London) (1995), 374, 327–337].

```

loop_
  _phasing_MAD_ratio.expt_id
  _phasing_MAD_ratio.clust_id
  _phasing_MAD_ratio.wavelength_1
  _phasing_MAD_ratio.wavelength_2
  _phasing_MAD_ratio.d_res_low
  _phasing_MAD_ratio.d_res_high
  _phasing_MAD_ratio.ratio_two_wl
  _phasing_MAD_ratio.ratio_one_wl
  _phasing_MAD_ratio.ratio_one_wl_centric
  1 '4 wavelength' 1.4013 1.4013 20.00 4.00 . 0.084 0.076
  1 '4 wavelength' 1.4013 1.3857 20.00 4.00 0.067 . .
  1 '4 wavelength' 1.4013 1.3852 20.00 4.00 0.051 . .
  1 '4 wavelength' 1.4013 1.3847 20.00 4.00 0.044 . .
  1 '4 wavelength' 1.3857 1.3857 20.00 4.00 . 0.110 0.049
  1 '4 wavelength' 1.3857 1.3852 20.00 4.00 0.049 . .
  1 '4 wavelength' 1.3857 1.3847 20.00 4.00 0.067 . .
  1 '4 wavelength' 1.3852 1.3852 20.00 4.00 . 0.149 0.072
  1 '4 wavelength' 1.3852 1.3847 20.00 4.00 0.039 . .
  1 '4 wavelength' 1.3847 1.3847 20.00 4.00 . 0.102 0.071

  1 '4 wavelength' 1.4013 1.4013 4.00 3.00 . 0.114 0.111
  1 '4 wavelength' 1.4013 1.3857 4.00 3.00 0.089 . .
  1 '4 wavelength' 1.4013 1.3852 4.00 3.00 0.086 . .
  1 '4 wavelength' 1.4013 1.3847 4.00 3.00 0.077 . .
  1 '4 wavelength' 1.3857 1.3857 4.00 3.00 . 0.140 0.127
  1 '4 wavelength' 1.3857 1.3852 4.00 3.00 0.085 . .
  1 '4 wavelength' 1.3857 1.3847 4.00 3.00 0.089 . .
  1 '4 wavelength' 1.3852 1.3852 4.00 3.00 . 0.155 0.119
  1 '4 wavelength' 1.3852 1.3847 4.00 3.00 0.082 . .
  1 '4 wavelength' 1.3847 1.3847 4.00 3.00 . 0.124 0.120

  1 '5 wavelength' 1.3857 1.3857 20.00 4.00 . 0.075 0.027
  1 '5 wavelength' 1.3857 1.3852 20.00 4.00 0.041 . .
  1 '5 wavelength' 1.3857 1.3847 20.00 4.00 0.060 . .
  1 '5 wavelength' 1.3857 1.3784 20.00 4.00 0.057 . .
  1 '5 wavelength' 1.3857 1.2862 20.00 4.00 0.072 . .
  1 '5 wavelength' 1.3852 1.3852 20.00 4.00 . 0.105 0.032
  1 '5 wavelength' 1.3852 1.3847 20.00 4.00 0.036 . .
  1 '5 wavelength' 1.3852 1.3784 20.00 4.00 0.044 . .
  1 '5 wavelength' 1.3852 1.2862 20.00 4.00 0.065 . .
  1 '5 wavelength' 1.3847 1.3847 20.00 4.00 . 0.072 0.031
  1 '5 wavelength' 1.3847 1.3784 20.00 4.00 0.040 . .
  1 '5 wavelength' 1.3847 1.2862 20.00 4.00 0.059 . .
  1 '5 wavelength' 1.3784 1.3784 20.00 4.00 . 0.059 0.032
  1 '5 wavelength' 1.3784 1.2862 20.00 4.00 0.059 . .
  1 '5 wavelength' 1.2862 1.3847 20.00 4.00 . 0.058 0.028

  1 '5 wavelength' 1.3857 1.3857 4.00 3.00 . 0.078 0.075
  1 '5 wavelength' 1.3857 1.3852 4.00 3.00 0.059 . .
  1 '5 wavelength' 1.3857 1.3847 4.00 3.00 0.067 . .
  1 '5 wavelength' 1.3857 1.3784 4.00 3.00 0.084 . .
  1 '5 wavelength' 1.3857 1.2862 4.00 3.00 0.073 . .
  1 '5 wavelength' 1.3852 1.3852 4.00 3.00 . 0.101 0.088
  1 '5 wavelength' 1.3852 1.3847 4.00 3.00 0.066 . .
  1 '5 wavelength' 1.3852 1.3784 4.00 3.00 0.082 . .
  1 '5 wavelength' 1.3852 1.2862 4.00 3.00 0.085 . .
  1 '5 wavelength' 1.3847 1.3847 4.00 3.00 . 0.097 0.074
  1 '5 wavelength' 1.3847 1.3784 4.00 3.00 0.081 . .
  1 '5 wavelength' 1.3847 1.2862 4.00 3.00 0.085 . .
  1 '5 wavelength' 1.3784 1.3784 4.00 3.00 . 0.114 0.089
  1 '5 wavelength' 1.3784 1.2862 4.00 3.00 0.103 . .
  1 '5 wavelength' 1.2862 1.2862 4.00 3.00 . 0.062 0.060

  2 '5 wavelength' 0.7263 0.7263 15.00 3.00 . 0.035 0.026
  2 '5 wavelength' 0.7263 0.7251 15.00 3.00 0.028 . .
  2 '5 wavelength' 0.7263 0.7284 15.00 3.00 0.023 . .
  2 '5 wavelength' 0.7263 0.7246 15.00 3.00 0.025 . .
  2 '5 wavelength' 0.7263 0.7217 15.00 3.00 0.026 . .
  2 '5 wavelength' 0.7251 0.7251 15.00 3.00 . 0.060 0.026
  2 '5 wavelength' 0.7251 0.7284 15.00 3.00 0.029 . .
  2 '5 wavelength' 0.7251 0.7246 15.00 3.00 0.031 . .
  2 '5 wavelength' 0.7251 0.7217 15.00 3.00 0.035 . .
  2 '5 wavelength' 0.7284 0.7284 15.00 3.00 . 0.075 0.030
  2 '5 wavelength' 0.7284 0.7246 15.00 3.00 0.023 . .
  2 '5 wavelength' 0.7284 0.7217 15.00 3.00 0.027 . .
  2 '5 wavelength' 0.7246 0.7246 15.00 3.00 . 0.069 0.026
  2 '5 wavelength' 0.7246 0.7217 15.00 3.00 0.024 . .
  2 '5 wavelength' 0.7217 0.7284 15.00 3.00 . 0.060 0.028

```

```

  2 '5 wavelength' 0.7263 0.7263 3.00 1.90 . 0.060 0.050
  2 '5 wavelength' 0.7263 0.7251 3.00 1.90 0.056 . .
  2 '5 wavelength' 0.7263 0.7284 3.00 1.90 0.055 . .
  2 '5 wavelength' 0.7263 0.7246 3.00 1.90 0.053 . .
  2 '5 wavelength' 0.7263 0.7217 3.00 1.90 0.056 . .
  2 '5 wavelength' 0.7251 0.7251 3.00 1.90 . 0.089 0.050
  2 '5 wavelength' 0.7251 0.7284 3.00 1.90 0.054 . .
  2 '5 wavelength' 0.7251 0.7246 3.00 1.90 0.058 . .
  2 '5 wavelength' 0.7251 0.7217 3.00 1.90 0.063 . .
  2 '5 wavelength' 0.7284 0.7284 3.00 1.90 . 0.104 0.057
  2 '5 wavelength' 0.7284 0.7246 3.00 1.90 0.052 . .
  2 '5 wavelength' 0.7284 0.7217 3.00 1.90 0.057 . .
  2 '5 wavelength' 0.7246 0.7246 3.00 1.90 . 0.098 0.052
  2 '5 wavelength' 0.7246 0.7217 3.00 1.90 0.054 . .
  2 '5 wavelength' 0.7217 0.7284 3.00 1.90 . 0.089 0.060

```

\* phasing\_MAD\_ratio.clust\_id

This data item is a pointer to phasing\_MAD\_clust.id in the PHASING\_MAD\_CLUSTER category.

phasing\_MAD\_ratio.d\_res\_high

(float)

The lowest value for the interplanar spacings for the reflection data used for the comparison of Bijvoet differences. This is called the highest resolution.

[phasing\_MAD\_ratio]

phasing\_MAD\_ratio.d\_res\_low

(float)

The highest value for the interplanar spacings for the reflection data used for the comparison of Bijvoet differences. This is called the lowest resolution.

[phasing\_MAD\_ratio]

\* phasing\_MAD\_ratio.expt\_id

This data item is a pointer to phasing\_MAD\_expt.id in the PHASING\_MAD\_EXPT category.

phasing\_MAD\_ratio.ratio\_one\_wl

(float)

The root-mean-square Bijvoet difference at one wavelength for all reflections.

[phasing\_MAD\_ratio]

phasing\_MAD\_ratio.ratio\_one\_wl\_centric

(float)

The root-mean-square Bijvoet difference at one wavelength for centric reflections. This would be equal to zero for perfect data and thus serves as an estimate of the noise in the anomalous signals.

[phasing\_MAD\_ratio]

phasing\_MAD\_ratio.ratio\_two\_wl

(float)

The root-mean-square dispersive Bijvoet difference between two wavelengths for all reflections.

[phasing\_MAD\_ratio]

\* phasing\_MAD\_ratio.wavelength\_1

This data item is a pointer to phasing\_MAD\_set.wavelength in the PHASING\_MAD\_SET category.

\* phasing\_MAD\_ratio.wavelength\_2

This data item is a pointer to phasing\_MAD\_set.wavelength in the PHASING\_MAD\_SET category.

**PHASING\_MAD\_SET**

Data items in the PHASING\_MAD\_SET category record details about the individual data sets used in a MAD phasing experiment.

Category group(s): **inclusive\_group**

**phasing\_group**

Category key(s): **phasing\_MAD\_set.expt\_id**

**phasing\_MAD\_set.clust\_id**

**phasing\_MAD\_set.set\_id**

**phasing\_MAD\_set.wavelength**

*Example 1 – based on a paper by Shapiro et al. [Nature (London) (1995), 374, 327–337].*

```

loop_
  phasing_MAD_set.expt_id
  phasing_MAD_set.clust_id
  phasing_MAD_set.set_id
  phasing_MAD_set.wavelength
  phasing_MAD_set.wavelength_details
  phasing_MAD_set.d_res_low
  phasing_MAD_set.d_res_high
  phasing_MAD_set.f_prime
  phasing_MAD_set.f_double_prime
1 '4 wavelength' aa 1.4013 'pre-edge'      20.00  3.00
  -12.48  3.80
1 '4 wavelength' bb 1.3857 'peak'          20.00  3.00
  -31.22 17.20
1 '4 wavelength' cc 1.3852 'edge'          20.00  3.00
  -13.97 29.17
1 '4 wavelength' dd 1.3847 'remote'        20.00  3.00
  -6.67 17.34
1 '5 wavelength' ee 1.3857 'ascending edge' 20.00  3.00
  -28.33 14.84
1 '5 wavelength' ff 1.3852 'peak'          20.00  3.00
  -21.50 30.23
1 '5 wavelength' gg 1.3847 'descending edge' 20.00  3.00
  -10.71 20.35
1 '5 wavelength' hh 1.3784 'remote 1'      20.00  3.00
  -14.45 11.84
1 '5 wavelength' ii 1.2862 'remote 2'      20.00  3.00
  -9.03  9.01
2 '5 wavelength' jj 0.7263 'pre-edge'      15.00  1.90
  -21.10  4.08
2 '5 wavelength' kk 0.7251 'edge'          15.00  1.90
  -34.72  7.92
2 '5 wavelength' ll 0.7248 'peak'          15.00  1.90
  -24.87 10.30
2 '5 wavelength' mm 0.7246 'descending edge' 15.00  1.90
  -17.43  9.62
2 '5 wavelength' nn 0.7217 'remote'        15.00  1.90
  -13.26  8.40

```

\* **phasing\_MAD\_set.clust\_id**

This data item is a pointer to **phasing\_MAD\_clust.id** in the PHASING\_MAD\_CLUST category.

**phasing\_MAD\_set.d\_res\_high** (float)

The lowest value for the interplanar spacings for the reflection data used for this set of data. This is called the highest resolution.

[phasing\_MAD\_set]

**phasing\_MAD\_set.d\_res\_low** (float)

The highest value for the interplanar spacings for the reflection data used for this set of data. This is called the lowest resolution.

[phasing\_MAD\_set]

\* **phasing\_MAD\_set.expt\_id**

This data item is a pointer to **phasing\_MAD\_expt.id** in the PHASING\_MAD\_EXPT category.

**phasing\_MAD\_set.f\_double\_prime** (float)

The  $f''$  component of the anomalous scattering factor for this wavelength.

[phasing\_MAD\_set]

**phasing\_MAD\_set.f\_prime** (float)

The  $f'$  component of the anomalous scattering factor for this wavelength.

[phasing\_MAD\_set]

\* **phasing\_MAD\_set.set\_id**

This data item is a pointer to **phasing\_set.id** in the PHASING\_SET category.

\* **phasing\_MAD\_set.wavelength** (float)

The wavelength at which this data set was measured.

The following item(s) have an equivalent role in their respective categories:

**phasing\_MAD\_ratio.wavelength\_1**,

**phasing\_MAD\_ratio.wavelength\_2**. [phasing\_MAD\_set]

**phasing\_MAD\_set.wavelength\_details** (text)

A descriptor for this wavelength in this cluster of data sets.

Examples: 'peak', 'remote', 'ascending edge'. [phasing\_MAD\_set]

**PHASING\_MIR**

Data items in the PHASING\_MIR category record details about the phasing of the structure where methods involving isomorphous replacement are involved. All isomorphous-replacement-based techniques are covered by this category, including single isomorphous replacement (SIR), multiple isomorphous replacement (MIR) and single or multiple isomorphous replacement plus anomalous scattering (SIRAS, MIRAS).

Category group(s): **inclusive\_group**

**phasing\_group**

Category key(s): **phasing\_MIR.entry\_id**

*Example 1 – based on a paper by Zanotti et al. [J. Biol. Chem. (1993), 268, 10728–10738].*

```

  phasing_MIR.method
; Standard phase refinement (Blow & Crick, 1959)
;

```

\* **phasing\_MIR.d\_res\_high** (float)

**phasing\_MIR.ebi\_d\_res\_high** (*ebi\_extensions 1.0*)

The lowest value in ångströms for the interplanar spacings for the reflection data used for the native data set. This is called the highest resolution.

The permitted range is [0.0, ∞).

[phasing\_MIR]

\* **phasing\_MIR.d\_res\_low** (float)

**phasing\_MIR.ebi\_d\_res\_low** (*ebi\_extensions 1.0*)

The highest value in ångströms for the interplanar spacings for the reflection data used for the native data set. This is called the lowest resolution.

The permitted range is [0.0, ∞).

[phasing\_MIR]

**phasing\_MIR.details** (text)

A description of special aspects of the isomorphous-replacement phasing.

[phasing\_MIR]

\* **phasing\_MIR.entry\_id**

This data item is a pointer to **entry.id** in the ENTRY category.

**phasing\_MIR.FOM** (float)  
phasing\_MIR.ebi\_fom (ebi\_extensions 1.0)

The mean value of the figure of merit  $m$  for all reflections phased in the native data set.

$$m = \frac{\int P_\alpha \exp(i\alpha) d\alpha}{\int P_\alpha d\alpha},$$

where  $P_\alpha$  = the probability that the phase angle  $\alpha$  is correct and the integral is taken over the range  $\alpha = 0$  to  $2\pi$ .

The permitted range is [0.0,  $\infty$ ). [phasing\_MIR]

**phasing\_MIR.FOM\_acentric** (float)  
phasing\_MIR.ebi\_fom\_acentric (ebi\_extensions 1.0)

The mean value of the figure of merit  $m$  for the acentric reflections phased in the native data set.

$$m = \frac{\int P_\alpha \exp(i\alpha) d\alpha}{\int P_\alpha d\alpha},$$

where  $P_\alpha$  = the probability that the phase angle  $\alpha$  is correct and the integral is taken over the range  $\alpha = 0$  to  $2\pi$ .

The permitted range is [0.0,  $\infty$ ). [phasing\_MIR]

**phasing\_MIR.FOM\_centric** (float)  
phasing\_MIR.ebi\_fom\_centric (ebi\_extensions 1.0)

The mean value of the figure of merit  $m$  for the centric reflections phased in the native data set.

$$m = \frac{\int P_\alpha \exp(i\alpha) d\alpha}{\int P_\alpha d\alpha},$$

where  $P_\alpha$  = the probability that the phase angle  $\alpha$  is correct and the integral is taken over the range  $\alpha = 0$  to  $2\pi$ .

The permitted range is [0.0,  $\infty$ ). [phasing\_MIR]

**phasing\_MIR.method** (text)

A description of the MIR phasing method applied to phase this structure. Note that this is not the computer program used, which is described in the SOFTWARE category, but rather the method itself. This data item should be used to describe significant methodological options used within the MIR phasing program.

[phasing\_MIR]

**phasing\_MIR.reflns** (int)

phasing\_MIR.ebi\_reflns (ebi\_extensions 1.0)

The total number of reflections phased in the native data set.

The permitted range is [0,  $\infty$ ). [phasing\_MIR]

**phasing\_MIR.reflns\_acentric** (int)

phasing\_MIR.ebi\_reflns\_acentric (ebi\_extensions 1.0)

The number of acentric reflections phased in the native data set.

The permitted range is [0,  $\infty$ ). [phasing\_MIR]

**phasing\_MIR.reflns\_centric** (int)

phasing\_MIR.ebi\_reflns\_centric (ebi\_extensions 1.0)

The number of centric reflections phased in the native data set.

The permitted range is [0,  $\infty$ ). [phasing\_MIR]

**phasing\_MIR.reflns\_criterion** (text)

phasing\_MIR.ebi\_reflns\_criteria (ebi\_extensions 1.0)

Criterion used to limit the reflections used in the phasing calculations.

Example: '> 4 \s(I)'. [phasing\_MIR]

### PHASING\_MIR\_DER

Data items in the PHASING\_MIR\_DER category record details about individual derivatives used in the phasing of the structure when methods involving isomorphous replacement are involved. A derivative in this context does not necessarily equate with a data set; for instance, the same data set could be used to one resolution limit as an isomorphous scatterer and to a different resolution (and with a different  $\sigma$  cut-off) as an anomalous scatterer. These would be treated as two distinct derivatives, although both derivatives would point to the same data sets via phasing\_MIR\_der.der\_set\_id and phasing\_MIR\_der.native\_set\_id.

Category group(s): inclusive\_group  
phasing\_group

Category key(s): phasing\_MIR\_der.id

Example 1 – based on a paper by Zanotti et al. [J. Biol. Chem. (1993), 268, 10728–10738].

loop_	<u>phasing_MIR_der.id</u>	<u>phasing_MIR_der.number_of_sites</u>	<u>phasing_MIR_der.details</u>
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'

\* **phasing\_MIR\_der.d\_res\_high** (float)

The lowest value for the interplanar spacings for the reflection data used for this derivative. This is called the highest resolution.

The permitted range is [0.0,  $\infty$ ). [phasing\_MIR\_der]

\* **phasing\_MIR\_der.d\_res\_low** (float)

The highest value for the interplanar spacings for the reflection data used for this derivative. This is called the lowest resolution.

The permitted range is [0.0,  $\infty$ ). [phasing\_MIR\_der]

\* **phasing\_MIR\_der.der\_set\_id**

The data set that was treated as the derivative in this experiment. This data item is a pointer to phasing\_set.id in the PHASING\_SET category.

**phasing\_MIR\_der.details** (text)

A description of special aspects of this derivative, its data, its solution or its use in phasing.

[phasing\_MIR\_der]

\* **phasing\_MIR\_der.id** (line)

The value of phasing\_MIR\_der.id must uniquely identify a record in the PHASING\_MIR\_DER list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

phasing\_MIR\_der.refln.der\_id,

phasing\_MIR\_der.shell.der\_id,

phasing\_MIR\_der.site.der\_id.

Examples: 'KAu(CN)2', 'K2HgI4.anom', 'K2HgI4.iso'. [phasing\_MIR\_der]

\* **phasing\_MIR\_der.native\_set\_id**

The data set that was treated as the native in this experiment. This data item is a pointer to phasing\_set.id in the PHASING\_SET category.

**phasing\_MIR\_der.number\_of\_sites** (int)

The number of heavy-atom sites in this derivative.

[phasing\_MIR\_der]

**phasing\_MIR\_der.power\_acentric** (float)  
phasing\_MIR\_der.ebi\_power\_acentric (ebi\_extensions 1.0)

The mean phasing power  $P$  for acentric reflections for this derivative.

$$P = \left( \frac{\sum |F(h)_{\text{calc}}|^2}{\sum |F(ph)_{\text{obs}} - F(ph)_{\text{calc}}|^2} \right)^{1/2},$$

where  $F(ph)_{\text{obs}}$  = the observed structure-factor amplitude of this derivative,  $F(ph)_{\text{calc}}$  = the calculated structure-factor amplitude of this derivative and  $F(h)_{\text{calc}}$  = the calculated structure-factor amplitude from the heavy-atom model; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞). [phasing\_MIR\_der]

**phasing\_MIR\_der.power\_centric** (float)  
phasing\_MIR\_der.ebi\_power\_centric (ebi\_extensions 1.0)

The mean phasing power  $P$  for centric reflections for this derivative.

$$P = \left( \frac{\sum |F(h)_{\text{calc}}|^2}{\sum |F(ph)_{\text{obs}} - F(ph)_{\text{calc}}|^2} \right)^{1/2},$$

where  $F(ph)_{\text{obs}}$  = the observed structure-factor amplitude of this derivative,  $F(ph)_{\text{calc}}$  = the calculated structure-factor amplitude of this derivative and  $F(h)_{\text{calc}}$  = the calculated structure-factor amplitude from the heavy-atom model; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞). [phasing\_MIR\_der]

**phasing\_MIR\_der.R\_cullis\_acentric** (float)  
phasing\_MIR\_der.ebi\_Rcullis\_acentric (ebi\_extensions 1.0)

Residual factor  $R_{\text{cullis,acen}}$  for acentric reflections for this derivative. The Cullis  $R$  factor was originally defined only for centric reflections. It is, however, also a useful statistical measure for acentric reflections, which is how it is used in this data item.

$$R_{\text{cullis,acen}} = \frac{\sum |F(ph)_{\text{obs}} \pm F(p)_{\text{obs}} - F(h)_{\text{calc}}|}{\sum |F(ph)_{\text{obs}} - F(p)_{\text{obs}}|},$$

where  $F(p)_{\text{obs}}$  = the observed structure-factor amplitude of the native,  $F(ph)_{\text{obs}}$  = the observed structure-factor amplitude of the derivative and  $F(h)_{\text{calc}}$  = the calculated structure-factor amplitude from the heavy-atom model; the sum is taken over the specified reflections.

Reference: Cullis, A. F., Muirhead, H., Perutz, M. F., Rossmann, M. G. & North, A. C. T. (1961). *Proc. R. Soc. London Ser. A*, **265**, 15–38.

The permitted range is [0.0, ∞). [phasing\_MIR\_der]

**phasing\_MIR\_der.R\_cullis\_anomalous** (float)  
phasing\_MIR\_der.ebi\_Rcullis\_anomalous (ebi\_extensions 1.0)

Residual factor  $R_{\text{cullis,ano}}$  for anomalous reflections for this derivative. The Cullis  $R$  factor was originally defined only for centric reflections. It is, however, also a useful statistical measure for anomalous reflections, which is how it is used in this data item. This is tabulated for acentric terms. A value less than 1.0 means there is some contribution to the phasing from the anomalous data.

$$R_{\text{cullis,ano}} = \frac{\sum |F(ph+)_{\text{obs}}F(ph-)_{\text{obs}} - F(h+)_{\text{calc}} - F(h-)_{\text{calc}}|}{\sum |F(ph+)_{\text{obs}} - F(ph-)_{\text{obs}}|},$$

where  $F(ph+)_{\text{obs}}$  = the observed positive Friedel structure-factor amplitude for the derivative,  $F(ph-)_{\text{obs}}$  = the observed negative Friedel structure-factor amplitude for the derivative,  $F(h+)_{\text{calc}}$  = the calculated positive Friedel structure-factor amplitude from the heavy-atom model and  $F(h-)_{\text{calc}}$  = the calculated negative Friedel structure-factor amplitude from the heavy-atom model; the sum is taken over the specified reflections.

Reference: Cullis, A. F., Muirhead, H., Perutz, M. F., Rossmann, M. G. & North, A. C. T. (1961). *Proc. R. Soc. London Ser. A*, **265**, 15–38.

The permitted range is [0.0, ∞). [phasing\_MIR\_der]

**phasing\_MIR\_der.R\_cullis\_centric** (float)  
phasing\_MIR\_der.ebi\_Rcullis\_centric (ebi\_extensions 1.0)

Residual factor  $R_{\text{cullis}}$  for centric reflections for this derivative.

$$R_{\text{cullis}} = \frac{\sum |F(ph)_{\text{obs}} \pm F(p)_{\text{obs}} - F(h)_{\text{calc}}|}{\sum |F(ph)_{\text{obs}} - F(p)_{\text{obs}}|},$$

where  $F(p)_{\text{obs}}$  = the observed structure-factor amplitude of the native,  $F(ph)_{\text{obs}}$  = the observed structure-factor amplitude of the derivative and  $F(h)_{\text{calc}}$  = the calculated structure-factor amplitude from the heavy-atom model; the sum is taken over the specified reflections.

Reference: Cullis, A. F., Muirhead, H., Perutz, M. F., Rossmann, M. G. & North, A. C. T. (1961). *Proc. R. Soc. London Ser. A*, **265**, 15–38.

The permitted range is [0.0, ∞). [phasing\_MIR\_der]

**phasing\_MIR\_der.reflns\_acentric** (int)  
phasing\_MIR\_der.ebi\_reflns\_acentric (ebi\_extensions 1.0)

The number of acentric reflections used in phasing for this derivative.

The permitted range is [0, ∞). [phasing\_MIR\_der]

**phasing\_MIR\_der.reflns\_anomalous** (int)  
phasing\_MIR\_der.ebi\_reflns\_anomalous (ebi\_extensions 1.0)

The number of anomalous reflections used in phasing for this derivative.

The permitted range is [0, ∞). [phasing\_MIR\_der]

**phasing\_MIR\_der.reflns\_centric** (int)  
phasing\_MIR\_der.ebi\_reflns\_centric (ebi\_extensions 1.0)

The number of centric reflections used in phasing for this derivative.

The permitted range is [0, ∞). [phasing\_MIR\_der]

**phasing\_MIR\_der.reflns\_criteria** (text)

Criteria used to limit the reflections used in the phasing calculations.

Example: '> 4 \s(I)'. [phasing\_MIR\_der]

## PHASING\_MIR\_DER\_REFLN

Data items in the PHASING\_MIR\_DER\_REFLN category record details about the calculated structure factors obtained in an MIR phasing experiment. This list may contain information from a number of different derivatives; `_phasing_MIR_der_refl_n.der_id` indicates to which derivative a given record corresponds. (A derivative in this context does not necessarily equate with a data set; see the definition of the PHASING\_MIR\_DER category for a discussion of the meaning of ‘derivative’.) It is not necessary for the data items describing the measured value of  $F$  to appear in this list, as they will be given in the PHASING\_SET\_REFLN category. However, these items can also be listed here for completeness.

Category group(s): `inclusive_group`  
`phasing_group`

Category key(s): `_phasing_MIR_der_refl_n.index_h`  
`_phasing_MIR_der_refl_n.index_k`  
`_phasing_MIR_der_refl_n.index_l`  
`_phasing_MIR_der_refl_n.der_id`  
`_phasing_MIR_der_refl_n.set_id`

Example 1 – based on laboratory records for the  $6r1,25$  reflection of an Hg/Pt derivative of protein NS1.

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

\* `_phasing_MIR_der_refl_n.der_id`

This data item is a pointer to `_phasing_MIR_der.id` in the PHASING\_MIR\_DER category.

`_phasing_MIR_der_refl_n.F_calc` (float)

The calculated value of the structure factor for this derivative, in electrons.

Related item: `_phasing_MIR_der_refl_n.F_calc_au` (conversion arbitrary).  
[phasing\_MIR\_der\_refl\_n]

`_phasing_MIR_der_refl_n.F_calc_au` (float)

The calculated value of the structure factor for this derivative, in arbitrary units.

Related item: `_phasing_MIR_der_refl_n.F_calc` (conversion arbitrary).  
[phasing\_MIR\_der\_refl\_n]

`_phasing_MIR_der_refl_n.F_meas` (float, su)

The measured value of the structure factor for this derivative, in electrons.

Related items: `_phasing_MIR_der_refl_n.F_meas_sigma` (associated esd),  
`_phasing_MIR_der_refl_n.F_meas_au` (conversion arbitrary).  
[phasing\_MIR\_der\_refl\_n]

`_phasing_MIR_der_refl_n.F_meas_au` (float, su)

The measured value of the structure factor for this derivative, in arbitrary units.

Related items: `_phasing_MIR_der_refl_n.F_meas_sigma_au` (associated esd),  
`_phasing_MIR_der_refl_n.F_meas` (conversion arbitrary).  
[phasing\_MIR\_der\_refl\_n]

`_phasing_MIR_der_refl_n.F_meas_sigma` (float)

The standard uncertainty (estimated standard deviation) of `_phasing_MIR_der_refl_n.F_meas`, in electrons.

Related items: `_phasing_MIR_der_refl_n.F_meas` (associated value),  
`_phasing_MIR_der_refl_n.F_meas_sigma_au` (conversion arbitrary).  
[phasing\_MIR\_der\_refl\_n]

`_phasing_MIR_der_refl_n.F_meas_sigma_au` (float)

The standard uncertainty (estimated standard deviation) of `_phasing_MIR_der_refl_n.F_meas_au`, in arbitrary units.

Related items: `_phasing_MIR_der_refl_n.F_meas_au` (associated value),  
`_phasing_MIR_der_refl_n.F_meas_sigma` (conversion arbitrary).  
[phasing\_MIR\_der\_refl\_n]

`_phasing_MIR_der_refl_n.HL_A_iso` (float)

The isomorphous Hendrickson–Lattman coefficient  $A_{iso}$  for this reflection for this derivative.

$$A_{iso} = -\frac{2.0[F(p)_{obs}^2 + F(h)_{calc}^2 - F(ph)_{obs}^2]F(p)_{obs} \cos(\alpha_{hcalc})}{E^2},$$

where  $E = [F(ph)_{obs} - F(p)_{obs} - F(h)_{calc}]^2$  for centric reflections,  $E = \{2^{1/2}[F(ph)_{obs} - F(p)_{obs}] - F(h)_{calc}\}^2$  for acentric reflections,  $F(p)_{obs}$  = the observed structure-factor amplitude of the native,  $F(ph)_{obs}$  = the observed structure-factor amplitude of the derivative,  $F(h)_{calc}$  = the calculated structure-factor amplitude from the heavy-atom model and  $\alpha_{hcalc}$  = the calculated phase from the heavy-atom model.

This coefficient appears in the expression for the phase probability of each isomorphous derivative:

$$P_i(\alpha) = \exp(k + A \cos \alpha + B \sin \alpha + C \cos 2\alpha + D \sin 2\alpha).$$

Reference: Hendrickson, W. A. & Lattman, E. E. (1970). *Acta Cryst.* B26, 136–143.

[phasing\_MIR\_der\_refl\_n]

`_phasing_MIR_der_refl_n.HL_B_iso` (float)

The isomorphous Hendrickson–Lattman coefficient  $B_{iso}$  for this reflection for this derivative.

$$B_{iso} = -\frac{2.0[F(p)_{obs}^2 + F(h)_{calc}^2 - F(ph)_{obs}^2]F(p)_{obs} \sin(\alpha_{hcalc})}{E^2},$$

where  $E = [F(ph)_{obs} - F(p)_{obs} - F(h)_{calc}]^2$  for centric reflections,  $E = \{2^{1/2}[F(ph)_{obs} - F(p)_{obs}] - F(h)_{calc}\}^2$  for acentric reflections,  $F(p)_{obs}$  = the observed structure-factor amplitude of the native,  $F(ph)_{obs}$  = the observed structure-factor amplitude of the derivative,  $F(h)_{calc}$  = the calculated structure-factor amplitude from the heavy-atom model and  $\alpha_{hcalc}$  = the calculated phase from the heavy-atom model.

This coefficient appears in the expression for the phase probability of each isomorphous derivative:

$$P_i(\alpha) = \exp(k + A \cos \alpha + B \sin \alpha + C \cos 2\alpha + D \sin 2\alpha).$$

Reference: Hendrickson, W. A. & Lattman, E. E. (1970). *Acta Cryst.* B26, 136–143.

[phasing\_MIR\_der\_refl\_n]

phasing\_MIR\_der\_refl.hl\_c\_iso (float)  
The isomorphous Hendrickson–Lattman coefficient  $C_{iso}$  for this reflection for this derivative.

$$C_{iso} = -\frac{F(p)_{obs}^2 [\sin(\alpha_{hcalc})^2 - \cos(\alpha_{hcalc})^2]}{E^2},$$

where  $E = [F(ph)_{obs} - F(p)_{obs} - F(h)_{calc}]^2$  for centric reflections,  $E = \{2^{1/2}[F(ph)_{obs} - F(p)_{obs}] - F(h)_{calc}\}^2$  for acentric reflections,  $F(p)_{obs}$  = the observed structure-factor amplitude of the native,  $F(ph)_{obs}$  = the observed structure-factor amplitude of the derivative,  $F(h)_{calc}$  = the calculated structure-factor amplitude from the heavy-atom model and  $\alpha_{hcalc}$  = the calculated phase from the heavy-atom model.

This coefficient appears in the expression for the phase probability of each isomorphous derivative:

$$P_i(\alpha) = \exp(k + A \cos \alpha + B \sin \alpha + C \cos 2\alpha + D \sin 2\alpha).$$

Reference: Hendrickson, W. A. & Lattman, E. E. (1970). *Acta Cryst.* B26, 136–143.

[phasing\_MIR\_der\_refl]

phasing\_MIR\_der\_refl.hl\_d\_iso (float)  
The isomorphous Hendrickson–Lattman coefficient  $D_{iso}$  for this reflection for this derivative.

$$D_{iso} = -\frac{2.0F(p)_{obs}^2 \sin(\alpha_{hcalc})^2 + \cos(\alpha_{hcalc})^2}{E^2},$$

where  $E = [F(ph)_{obs} - F(p)_{obs} - F(h)_{calc}]^2$  for centric reflections,  $E = \{2^{1/2}[F(ph)_{obs} - F(p)_{obs}] - F(h)_{calc}\}^2$  for acentric reflections,  $F(p)_{obs}$  = the observed structure-factor amplitude of the native,  $F(ph)_{obs}$  = the observed structure-factor amplitude of the derivative,  $F(h)_{calc}$  = the calculated structure-factor amplitude from the heavy-atom model and  $\alpha_{hcalc}$  = the calculated phase from the heavy-atom model.

This coefficient appears in the expression for the phase probability of each isomorphous derivative:

$$P_i(\alpha) = \exp(k + A \cos \alpha + B \sin \alpha + C \cos 2\alpha + D \sin 2\alpha).$$

Reference: Hendrickson, W. A. & Lattman, E. E. (1970). *Acta Cryst.* B26, 136–143.

[phasing\_MIR\_der\_refl]

\* phasing\_MIR\_der\_refl.index\_h (int)  
Miller index  $h$  for this reflection for this derivative.

[phasing\_MIR\_der\_refl]

\* phasing\_MIR\_der\_refl.index\_k (int)  
Miller index  $k$  for this reflection for this derivative.

[phasing\_MIR\_der\_refl]

\* phasing\_MIR\_der\_refl.index\_l (int)  
Miller index  $l$  for this reflection for this derivative.

[phasing\_MIR\_der\_refl]

phasing\_MIR\_der\_refl.phase\_calc (float)  
The calculated value of the structure-factor phase based on the heavy-atom model for this derivative in degrees.

[phasing\_MIR\_der\_refl]

\* phasing\_MIR\_der\_refl.set\_id  
This data item is a pointer to phasing\_set.id in the PHASING\_SET category.

**PHASING\_MIR\_DER\_SHELL**

Data items in the PHASING\_MIR\_DER\_SHELL category record statistics, broken down into shells of resolution, for an MIR phasing experiment. This list may contain information from a number of different derivatives; phasing\_MIR\_der\_shell.der\_id indicates to which derivative a given record corresponds. (A derivative in this context does not necessarily equate with a data set; see the definition of the PHASING\_MIR\_DER category for a discussion of the meaning of ‘derivative’.)

Category group(s): inclusive\_group

phasing\_group

Category key(s): phasing\_MIR\_der\_shell.der\_id

phasing\_MIR\_der\_shell.d\_res\_low

phasing\_MIR\_der\_shell.d\_res\_high

Example 1 – based on a paper by Zanotti et al. [*J. Biol. Chem.* (1993), 268, 10728–10738] with addition of an arbitrary low-resolution limit.

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

KAu (CN) 2 6.4 5.2 50 20

KAu (CN) 2 5.2 4.4 44 23

KAu (CN) 2 4.4 3.8 39 23

KAu (CN) 2 3.8 3.4 33 21

KAu (CN) 2 3.4 3.0 28 17

KAu (CN) 2 15.0 3.0 38 21

K2HgI4 15.0 8.3 149 87

K2HgI4 8.3 6.4 121 73

K2HgI4 6.4 5.2 95 61

K2HgI4 5.2 4.4 80 60

K2HgI4 4.4 3.8 73 63

K2HgI4 3.8 3.4 68 57

K2HgI4 3.4 3.0 63 46

K2HgI4 15.0 3.0 79 58

K3IrCl6 15.0 8.3 33 27

K3IrCl6 8.3 6.4 40 23

K3IrCl6 6.4 5.2 31 22

K3IrCl6 5.2 4.4 27 23

K3IrCl6 4.4 3.8 22 23

K3IrCl6 3.8 3.4 19 20

K3IrCl6 3.4 3.0 16 20

K3IrCl6 15.0 3.0 23 21

\* phasing\_MIR\_der\_shell.d\_res\_high (float)

The lowest value for the interplanar spacings for the reflection data for this derivative in this shell. This is called the highest resolution.

The permitted range is [0.0, ∞).

[phasing\_MIR\_der\_shell]

\* phasing\_MIR\_der\_shell.d\_res\_low (float)

The highest value for the interplanar spacings for the reflection data for this derivative in this shell. This is called the lowest resolution.

The permitted range is [0.0, ∞).

[phasing\_MIR\_der\_shell]

\* phasing\_MIR\_der\_shell.der\_id

This data item is a pointer to phasing\_MIR\_der.id in the PHASING\_MIR\_DER category.

phasing\_MIR\_der\_shell.fom (float)

The mean value of the figure of merit  $m$  for reflections in this shell.

$$m = \frac{\int P_{\alpha} \exp(i\alpha) d\alpha}{\int P_{\alpha} d\alpha},$$

where  $P_{\alpha}$  = the probability that the phase angle  $\alpha$  is correct; the integral is taken over the range  $\alpha = 0$  to  $2\pi$ .

The permitted range is [0.0, ∞).

[phasing\_MIR\_der\_shell]

phasing\_MIR\_der\_shell.ha\_ampl (float)  
The mean heavy-atom amplitude for reflections for this derivative in this shell.  
The permitted range is [0.0, ∞). [phasing\_MIR\_der\_shell]

phasing\_MIR\_der\_shell.loc (float)  
The mean lack-of-closure error loc for reflections for this derivative in this shell.

$$\text{loc} = \sum |F(ph)_{\text{obs}} - F(ph)_{\text{calc}}|,$$

where  $F(ph)_{\text{obs}}$  = the observed structure-factor amplitude of the derivative and  $F(ph)_{\text{calc}}$  = the calculated structure-factor amplitude of the derivative; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞). [phasing\_MIR\_der\_shell]

phasing\_MIR\_der\_shell.phase (float)  
The mean of the phase values for reflections for this derivative in this shell.

[phasing\_MIR\_der\_shell]

phasing\_MIR\_der\_shell.power (float)  
The mean phasing power  $P$  for reflections for this derivative in this shell.

$$P = \left( \frac{\sum |F(h)_{\text{calc}}|^2}{\sum |F(ph)_{\text{obs}} - F(ph)_{\text{calc}}|^2} \right)^{1/2},$$

where  $F(ph)_{\text{obs}}$  = the observed structure-factor amplitude of this derivative,  $F(ph)_{\text{calc}}$  = the calculated structure-factor amplitude of this derivative and  $F(h)_{\text{calc}}$  = the calculated structure-factor amplitude from the heavy-atom model; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞). [phasing\_MIR\_der\_shell]

phasing\_MIR\_der\_shell.R\_cullis (float)  
Residual factor  $R_{\text{cullis}}$  for centric reflections for this derivative in this shell.

$$R_{\text{cullis}} = \frac{\sum |F(ph)_{\text{obs}} \pm F(p)_{\text{obs}} - F(h)_{\text{calc}}|}{\sum |F(ph)_{\text{obs}} - F(p)_{\text{obs}}|},$$

where  $F(p)_{\text{obs}}$  = the observed structure-factor amplitude of the native,  $F(ph)_{\text{obs}}$  = the observed structure-factor amplitude of the derivative and  $F(h)_{\text{calc}}$  = the calculated structure-factor amplitude from the heavy-atom model; the sum is taken over the specified reflections.

Reference: Cullis, A. F., Muirhead, H., Perutz, M. F., Rossmann, M. G. & North, A. C. T. (1961). *Proc. R. Soc. London Ser. A*, **265**, 15–38.

The permitted range is [0.0, ∞). [phasing\_MIR\_der\_shell]

phasing\_MIR\_der\_shell.R\_kraut (float)  
Residual factor  $R_{\text{kraut}}$  for general reflections for this derivative in this shell.

$$R_{\text{kraut}} = \frac{\sum |F(ph)_{\text{obs}} - F(ph)_{\text{calc}}|}{\sum |F(ph)_{\text{obs}}|},$$

where  $F(ph)_{\text{obs}}$  = the observed structure-factor amplitude of the derivative,  $F(ph)_{\text{calc}}$  = the calculated structure-factor amplitude of the derivative and the sum is taken over the specified reflections.

Reference: Kraut, J., Sieker, L. C., High, D. F. & Freer, S. T. (1962). *Proc. Natl Acad. Sci. USA*, **48**, 1417–1424.

The permitted range is [0.0, ∞). [phasing\_MIR\_der\_shell]

phasing\_MIR\_der\_shell.reflns (int)  
The number of reflections in this shell.  
The permitted range is [0, ∞). [phasing\_MIR\_der\_shell]

### PHASING\_MIR\_DER\_SITE

Data items in the PHASING\_MIR\_DER\_SITE category record details about the heavy-atom sites in an MIR phasing experiment. This list may contain information from a number of different derivatives; phasing\_MIR\_der\_site.der\_id indicates to which derivative a given record corresponds. (A derivative in this context does not necessarily equate with a data set; see the definition of the PHASING\_MIR\_DER category for a discussion of the meaning of ‘derivative’.)

Category group(s): inclusive\_group

phasing\_group

Category key(s): phasing\_MIR\_der\_site.der\_id

phasing\_MIR\_der\_site.id

Example 1 – based on a paper by Zanotti et al. [*J. Biol. Chem.* (1993), **268**, 10728–10738] with occupancies converted from electrons to fractional.

```
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
KAu (CN) 2 3 Au 0.02 0.263 0.782 0.906 15.7
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
K2HgI4 3 Hg 0.23 0.974 0.455 0.974 24.2
K2HgI4 4 Hg 0.28 0.903 0.836 0.859 14.7
K2HgI4 5 Hg 0.07 0.489 0.200 0.885 6.4
K2HgI4 6 Hg 0.07 0.162 0.799 0.889 32.9
K3IrCl6 1 Ir 0.26 0.209 0.739 0.758 40.8
K3IrCl6 2 Ir 0.05 0.279 0.613 0.752 24.9
```

\* phasing\_MIR\_der\_site.atom\_type\_symbol

This data item is a pointer to atom\_type.symbol in the ATOM\_TYPE category. The scattering factors referenced *via* this data item should be those used in the refinement of the heavy-atom data; in some cases this is the scattering factor for the single heavy atom, in other cases these are the scattering factors for an atomic cluster.

phasing\_MIR\_der\_site.B\_iso (float, su)  
Isotropic displacement parameter for this heavy-atom site in this derivative.

Related item: phasing\_MIR\_der\_site.B\_iso\_esd (associated esd).

[phasing\_MIR\_der\_site]

phasing\_MIR\_der\_site.B\_iso\_esd (float)  
The standard uncertainty (estimated standard deviation) of phasing\_MIR\_der\_site.B\_iso.

Related item: phasing\_MIR\_der\_site.B\_iso (associated value).

[phasing\_MIR\_der\_site]

phasing\_MIR\_der\_site.Cartn\_x (float, su)  
The  $x$  coordinate of this heavy-atom position in this derivative specified as orthogonal ångströms. The orthogonal Cartesian axes are related to the cell axes as specified by the description given in atom\_sites.Cartn\_transform\_axes.

Related item: phasing\_MIR\_der\_site.Cartn\_x\_esd (associated esd).

[phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.Cartn\_x\_esd** (float)  
 The standard uncertainty (estimated standard deviation) of phasing\_MIR\_der\_site.Cartn\_x.  
 Related item: phasing\_MIR\_der\_site.Cartn\_x (associated value).  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.Cartn\_y** (float, su)  
 The y coordinate of this heavy-atom position in this derivative specified as orthogonal ångströms. The orthogonal Cartesian axes are related to the cell axes as specified by the description given in atom\_sites.Cartn\_transform\_axes.  
 Related item: phasing\_MIR\_der\_site.Cartn\_y\_esd (associated esd).  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.Cartn\_y\_esd** (float)  
 The standard uncertainty (estimated standard deviation) of phasing\_MIR\_der\_site.Cartn\_y.  
 Related item: phasing\_MIR\_der\_site.Cartn\_y (associated value).  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.Cartn\_z** (float, su)  
 The z coordinate of this heavy-atom position in this derivative specified as orthogonal ångströms. The orthogonal Cartesian axes are related to the cell axes as specified by the description given in atom\_sites.Cartn\_transform\_axes.  
 Related item: phasing\_MIR\_der\_site.Cartn\_z\_esd (associated esd).  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.Cartn\_z\_esd** (float)  
 The standard uncertainty (estimated standard deviation) of phasing\_MIR\_der\_site.Cartn\_z.  
 Related item: phasing\_MIR\_der\_site.Cartn\_z (associated value).  
 [phasing\_MIR\_der\_site]

\* **phasing\_MIR\_der\_site.der\_id**  
 This data item is a pointer to phasing\_MIR\_der.id in the PHASING\_MIR\_DER category.

**phasing\_MIR\_der\_site.details** (text)  
 A description of special aspects of the derivative site.  
 Examples: 'binds to His 117',  
 'minor site obtained from difference Fourier',  
 'same as site 2 in the K2HgI4 derivative'.  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.fract\_x** (float, su)  
 The x coordinate of this heavy-atom position in this derivative specified as a fraction of cell.length\_a.  
 Related item: phasing\_MIR\_der\_site.fract\_x\_esd (associated esd).  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.fract\_x\_esd** (float)  
 The standard uncertainty (estimated standard deviation) of phasing\_MIR\_der\_site.fract\_x.  
 Related item: phasing\_MIR\_der\_site.fract\_x (associated value).  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.fract\_y** (float, su)  
 The y coordinate of this heavy-atom position in this derivative specified as a fraction of cell.length\_b.  
 Related item: phasing\_MIR\_der\_site.fract\_y\_esd (associated esd).  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.fract\_y\_esd** (float)  
 The standard uncertainty (estimated standard deviation) of phasing\_MIR\_der\_site.fract\_y.  
 Related item: phasing\_MIR\_der\_site.fract\_y (associated value).  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.fract\_z** (float, su)  
 The z coordinate of this heavy-atom position in this derivative specified as a fraction of cell.length\_c.  
 Related item: phasing\_MIR\_der\_site.fract\_z\_esd (associated esd).  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.fract\_z\_esd** (float)  
 The standard uncertainty (estimated standard deviation) of phasing\_MIR\_der\_site.fract\_z.  
 Related item: phasing\_MIR\_der\_site.fract\_z (associated value).  
 [phasing\_MIR\_der\_site]

\* **phasing\_MIR\_der\_site.id** (code)  
 The value of phasing\_MIR\_der\_site.id must uniquely identify each site in each derivative in the PHASING\_MIR\_DER\_SITE list. The atom identifiers need not be unique over all sites in all derivatives; they need only be unique for each site in each derivative. Note that this item need not be a number; it can be any unique identifier.  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.occupancy** (float)  
 The fraction of the atom type present at this heavy-atom site in a given derivative. The sum of the occupancies of all the atom types at this site may not significantly exceed 1.0 unless it is a dummy site.  
 The permitted range is [0.0, ∞). Where no value is given, the assumed value is '1.0'.  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.occupancy\_anom** (float, su)  
**phasing\_MIR\_der\_site.ebi\_occupancy\_anom** (ebi\_extensions 1.0)  
 The relative anomalous occupancy of the atom type present at this heavy-atom site in a given derivative. This atom occupancy will probably be on an arbitrary scale.  
 Related item: phasing\_MIR\_der\_site.occupancy\_anom\_su (associated esd).  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.occupancy\_anom\_su** (float)  
**phasing\_MIR\_der\_site.ebi\_occupancy\_anom\_esd** (ebi\_extensions 1.0)  
 The standard uncertainty (estimated standard deviation) of phasing\_MIR\_der\_site.occupancy\_anom.  
 Related item: phasing\_MIR\_der\_site.occupancy\_anom (associated value).  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.occupancy\_iso** (float, su)  
**phasing\_MIR\_der\_site.ebi\_occupancy\_iso** (ebi\_extensions 1.0)  
 The relative real isotropic occupancy of the atom type present at this heavy-atom site in a given derivative. This atom occupancy will probably be on an arbitrary scale.  
 Related item: phasing\_MIR\_der\_site.occupancy\_iso\_su (associated esd).  
 [phasing\_MIR\_der\_site]

**phasing\_MIR\_der\_site.occupancy\_iso\_su** (float)  
**phasing\_MIR\_der\_site.ebi\_occupancy\_iso\_esd** (ebi\_extensions 1.0)  
 The standard uncertainty (estimated standard deviation) of phasing\_MIR\_der\_site.occupancy\_iso.  
 Related item: phasing\_MIR\_der\_site.occupancy\_iso (associated value).  
 [phasing\_MIR\_der\_site]

## PHASING\_MIR\_SHELL

Data items in the PHASING\_MIR\_SHELL category record statistics for an isomorphous replacement phasing experiment broken down into shells of resolution.

Category group(s): **inclusive\_group**

**phasing\_group**

Category key(s): **\_phasing\_MIR\_shell.d\_res\_low**

**\_phasing\_MIR\_shell.d\_res\_high**

Example 1 – based on a paper by Zanotti et al. [*J. Biol. Chem.* (1993), **268**, 10728–10738] with addition of an arbitrary low-resolution limit.

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

\* **\_phasing\_MIR\_shell.d\_res\_high** (float)

The lowest value for the interplanar spacings for the reflection data in this shell. This is called the highest resolution. Note that the resolution limits of shells in the items **\_phasing\_MIR\_shell.d\_res\_high** and **\_phasing\_MIR\_shell.d\_res\_low** are independent of the resolution limits of shells in the items **\_reflns\_shell.d\_res\_high** and **\_reflns\_shell.d\_res\_low**.

The permitted range is [0.0, ∞).

[phasing\_MIR\_shell]

\* **\_phasing\_MIR\_shell.d\_res\_low** (float)

The highest value for the interplanar spacings for the reflection data in this shell. This is called the lowest resolution. Note that the resolution limits of shells in the items **\_phasing\_MIR\_shell.d\_res\_high** and **\_phasing\_MIR\_shell.d\_res\_low** are independent of the resolution limits of shells in the items **\_reflns\_shell.d\_res\_high** and **\_reflns\_shell.d\_res\_low**.

The permitted range is [0.0, ∞).

[phasing\_MIR\_shell]

**\_phasing\_MIR\_shell.FOM** (float)

The mean value of the figure of merit  $m$  for reflections in this shell.

$$m = \frac{\int P_{\alpha} \exp(i\alpha) d\alpha}{\int P_{\alpha} d\alpha},$$

where  $P_{\alpha}$  = the probability that the phase angle  $\alpha$  is correct and the integral is taken over the range  $\alpha = 0$  to  $2\pi$ .

The permitted range is [0.0, ∞).

[phasing\_MIR\_shell]

**\_phasing\_MIR\_shell.FOM\_acentric** (float)

**\_phasing\_MIR\_shell.ebi\_fom\_acentric** (ebi\_extensions 1.0)

The mean value of the figure of merit  $m$  for acentric reflections in this shell.

$$m = \frac{\int P_{\alpha} \exp(i\alpha) d\alpha}{\int P_{\alpha} d\alpha},$$

where  $P_{\alpha}$  = the probability that the phase angle  $\alpha$  is correct and the integral is taken over the range  $\alpha = 0$  to  $2\pi$ .

The permitted range is [0.0, ∞).

[phasing\_MIR\_shell]

**\_phasing\_MIR\_shell.FOM\_centric** (float)

**\_phasing\_MIR\_shell.ebi\_fom\_centric** (ebi\_extensions 1.0)

The mean value of the figure of merit  $m$  for centric reflections in this shell.

$$m = \frac{\int P_{\alpha} \exp(i\alpha) d\alpha}{\int P_{\alpha} d\alpha},$$

where  $P_{\alpha}$  = the probability that the phase angle  $\alpha$  is correct and the integral is taken over the range  $\alpha = 0$  to  $2\pi$ .

The permitted range is [0.0, ∞).

[phasing\_MIR\_shell]

**\_phasing\_MIR\_shell.loc** (float)

The mean lack-of-closure error loc for reflections in this shell.

$$\text{loc} = \sum |F(ph)_{\text{obs}} - F(ph)_{\text{calc}}|,$$

where  $F(ph)_{\text{obs}}$  = the observed structure-factor amplitude of the derivative and  $F(ph)_{\text{calc}}$  = the calculated structure-factor amplitude of the derivative; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

[phasing\_MIR\_shell]

**\_phasing\_MIR\_shell.mean\_phase** (float)

The mean of the phase values for all reflections in this shell.

[phasing\_MIR\_shell]

**\_phasing\_MIR\_shell.power** (float)

The mean phasing power  $P$  for reflections in this shell.

$$P = \left( \frac{\sum |F(h)_{\text{calc}}|^2}{\sum |F(ph)_{\text{obs}} - F(ph)_{\text{calc}}|^2} \right)^{1/2},$$

where  $F(ph)_{\text{obs}}$  = the observed structure-factor amplitude of this derivative,  $F(ph)_{\text{calc}}$  = the calculated structure-factor amplitude of this derivative and  $F(h)_{\text{calc}}$  = the calculated structure-factor amplitude from the heavy-atom model; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

[phasing\_MIR\_shell]

**\_phasing\_MIR\_shell.R\_cullis** (float)

Residual factor  $R_{\text{cullis}}$  for centric reflections in this shell.

$$R_{\text{cullis}} = \frac{\sum |F(ph)_{\text{obs}} \pm F(p)_{\text{obs}} - F(h)_{\text{calc}}|}{\sum |F(ph)_{\text{obs}} - F(p)_{\text{obs}}|},$$

where  $F(p)_{\text{obs}}$  = the observed structure-factor amplitude of the native,  $F(ph)_{\text{obs}}$  = the observed structure-factor amplitude of the derivative and  $F(h)_{\text{calc}}$  = the calculated structure-factor amplitude from the heavy-atom model; the sum is taken over the specified reflections.

Reference: Cullis, A. F., Muirhead, H., Perutz, M. F., Rossmann, M. G. & North, A. C. T. (1961). *Proc. R. Soc. London Ser. A*, **265**, 15–38.

The permitted range is [0.0, ∞).

[phasing\_MIR\_shell]

**\_phasing\_MIR\_shell.R\_kraut** (float)

Residual factor  $R_{\text{kraut}}$  for general reflections in this shell.

$$R_{\text{kraut}} = \frac{\sum |F(ph)_{\text{obs}} - F(ph)_{\text{calc}}|}{\sum |F(ph)_{\text{obs}}|},$$

where  $F(ph)_{\text{obs}}$  = the observed structure-factor amplitude of the derivative and  $F(ph)_{\text{calc}}$  = the calculated structure-factor amplitude of the derivative; the sum is taken over the specified reflections.

Reference: Kraut, J., Sieker, L. C., High, D. F. & Freer, S. T. (1962). *Proc. Natl Acad. Sci. USA*, **48**, 1417–1424.

The permitted range is [0.0, ∞).

[phasing\_MIR\_shell]

**phasing\_MIR\_shell.reflns** (int)  
 The number of reflections in this shell.  
 The permitted range is [0, ∞). [phasing\_MIR\_shell]

**phasing\_MIR\_shell.reflns\_acentric** (int)  
*phasing\_MIR\_shell.ebi\_reflns\_acentric(ebi\_extensions 1.0)*  
 The number of acentric reflections in this shell.  
 The permitted range is [0, ∞). [phasing\_MIR\_shell]

**phasing\_MIR\_shell.reflns\_anomalous** (int)  
 The number of anomalous reflections in this shell.  
 The permitted range is [0, ∞). [phasing\_MIR\_shell]

**phasing\_MIR\_shell.reflns\_centric** (int)  
*phasing\_MIR\_shell.ebi\_reflns\_centric(ebi\_extensions 1.0)*  
 The number of centric reflections in this shell.  
 The permitted range is [0, ∞). [phasing\_MIR\_shell]

**phasing\_set.cell\_length\_a** (float)  
 Unit-cell length *a* for this data set in ångströms.  
 The permitted range is [0.0, ∞). [phasing\_set]

**phasing\_set.cell\_length\_b** (float)  
 Unit-cell length *b* for this data set in ångströms.  
 The permitted range is [0.0, ∞). [phasing\_set]

**phasing\_set.cell\_length\_c** (float)  
 Unit-cell length *c* for this data set in ångströms.  
 The permitted range is [0.0, ∞). [phasing\_set]

**phasing\_set.detector\_specific** (text)  
 The particular radiation detector. In general, this will be a manufacturer, description, model number or some combination of these.  
 Examples: 'Siemens model x', 'Kodak XG', 'MAR Research model y'. [phasing\_set]

**phasing\_set.detector\_type** (text)  
 The general class of the radiation detector.  
 Examples: 'multiwire', 'imaging plate', 'CCD', 'film'. [phasing\_set]

\* **phasing\_set.id** (line)  
 The value of phasing\_set.id must uniquely identify a record in the PHASING\_SET list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

phasing\_set.refln.set\_id,  
phasing\_MAD\_set.set\_id,  
phasing\_MIR\_der.der\_set\_id,  
phasing\_MIR\_der.native\_set\_id,  
phasing\_MIR\_der.refln.set\_id.  
 Examples: 'KAu (CN) 2', 'K2HgI4'. [phasing\_set]

**phasing\_set.radiation\_source\_specific** (text)  
 The particular source of radiation. In general, this will be a manufacturer, description, or model number (or some combination of these) for laboratory sources and an institution name and beamline name for synchrotron sources.  
 Examples: 'Rigaku RU200', 'Philips fine focus Mo', 'NSLS beamline X8C'. [phasing\_set]

**phasing\_set.cell\_angle\_alpha** (float)  
 Unit-cell angle  $\alpha$  for this data set in degrees.  
 The permitted range is [0.0, 180.0]. Where no value is given, the assumed value is '90.0'. [phasing\_set]

**phasing\_set.cell\_angle\_beta** (float)  
 Unit-cell angle  $\beta$  for this data set in degrees.  
 The permitted range is [0.0, 180.0]. Where no value is given, the assumed value is '90.0'. [phasing\_set]

**phasing\_set.cell\_angle\_gamma** (float)  
 Unit-cell angle  $\gamma$  for this data set in degrees.  
 The permitted range is [0.0, 180.0]. Where no value is given, the assumed value is '90.0'. [phasing\_set]

**phasing\_set.radiation\_wavelength** (float)  
 The mean wavelength of the radiation used to measure this data set.  
 The permitted range is [0.0, ∞). [phasing\_set]

**phasing\_set.temp** (float)  
 The temperature in kelvins at which the data set was measured.  
 The permitted range is [0.0, ∞). [phasing\_set]

**PHASING\_SET**

Data items in the PHASING\_SET category record details about the data sets used in a phasing experiment. A given data set may be used in a number of different ways; for instance, a single data set could be used both as an isomorphous derivative and as a component of a multiple-wavelength calculation. This category establishes identifiers for each data set and permits the archiving of a subset of experimental information for each data set (cell constants, wavelength, temperature *etc.*). This and related categories of data items are provided so that derivative intensity and phase information can be stored in the same data block as the information for the refined structure. If all the possible experimental information for each data set (raw data sets, crystal growth conditions *etc.*) is to be archived, these data items should be recorded in a separate data block.

Category group(s): **inclusive\_group**  
                                   **phasing\_group**

Category key(s): **phasing\_set.id**

---

*Example 1 – based on laboratory records for an Hg/Pt derivative of protein NSI.*

<u>phasing_set.id</u>	'NS1-96'
<u>phasing_set.cell_angle_alpha</u>	90.0
<u>phasing_set.cell_angle_beta</u>	90.0
<u>phasing_set.cell_angle_gamma</u>	90.0
<u>phasing_set.cell_length_a</u>	38.63
<u>phasing_set.cell_length_b</u>	38.63
<u>phasing_set.cell_length_c</u>	82.88
<u>phasing_set.radiation_wavelength</u>	1.5145
<u>phasing_set.detector_type</u>	'image plate'
<u>phasing_set.detector_specific</u>	'RXII'

## PHASING\_SET\_REFLN

Data items in the PHASING\_SET\_REFLN category record the values of the measured structure factors used in a phasing experiment. This list may contain information from a number of different data sets; `_phasing_set_reflн.set_id` indicates the data set to which a given record corresponds.

Category group(s): `inclusive_group`  
`phasing_group`

Category key(s): `_phasing_set_reflн.index_h`  
`_phasing_set_reflн.index_k`  
`_phasing_set_reflн.index_l`  
`_phasing_set_reflн.set_id`

*Example 1 – based on laboratory records for the 15,15,32 reflection of an Hg/Pt derivative of protein NS1.*

```
_phasing_set_reflн.set_id      'NS1-96'
_phasing_set_reflн.index_h     15
_phasing_set_reflн.index_k     15
_phasing_set_reflн.index_l     32
_phasing_set_reflн.F_meas_au   181.79
_phasing_set_reflн.F_meas_sigma_au  3.72
```

`_phasing_set_reflн.F_meas` (float, su)

The measured value of the structure factor for this reflection in this data set in electrons.

Related items: `_phasing_set_reflн.F_meas_sigma` (associated esd),

`_phasing_set_reflн.F_meas_au` (conversion arbitrary).

[`phasing_set_reflн`]

`_phasing_set_reflн.F_meas_au` (float, su)

The measured value of the structure factor for this reflection in this data set in arbitrary units.

Related items: `_phasing_set_reflн.F_meas_sigma_au` (associated esd),

`_phasing_set_reflн.F_meas` (conversion arbitrary).

[`phasing_set_reflн`]

`_phasing_set_reflн.F_meas_sigma` (float)

The standard uncertainty (estimated standard deviation) of `_phasing_set_reflн.F_meas` in electrons.

Related items: `_phasing_set_reflн.F_meas` (associated value),

`_phasing_set_reflн.F_meas_sigma_au` (conversion arbitrary).

[`phasing_set_reflн`]

`_phasing_set_reflн.F_meas_sigma_au` (float)

The standard uncertainty (estimated standard deviation) of `_phasing_set_reflн.F_meas_au` in arbitrary units.

Related items: `_phasing_set_reflн.F_meas_au` (associated value),

`_phasing_set_reflн.F_meas_sigma` (conversion arbitrary).

[`phasing_set_reflн`]

\* `_phasing_set_reflн.index_h` (int)

Miller index *h* of this reflection in this data set.

[`phasing_set_reflн`]

\* `_phasing_set_reflн.index_k` (int)

Miller index *k* of this reflection in this data set.

[`phasing_set_reflн`]

\* `_phasing_set_reflн.index_l` (int)

Miller index *l* of this reflection in this data set.

[`phasing_set_reflн`]

\* `_phasing_set_reflн.set_id`

This data item is a pointer to `_phasing_set.id` in the PHASING\_SET category.

## PUBL

Data items in the PUBL category are used when submitting a manuscript for publication.

Category group(s): `inclusive_group`  
`iucr_group`

Category key(s): `_publ.entry_id`

*Example 1 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].*

```
_publ.section_title
; trans-3-Benzoyl-2-(tert-butyl)-4-(iso-butyl)-
  1,3-oxazolidin-5-one
;
```

```
_publ.section_abstract
```

```
; The oxazolidinone ring is a shallow envelope
conformation with the tert-butyl and iso-butyl groups
occupying trans-positions with respect to the ring. The
angles at the N atom sum to 356.2%, indicating a very
small degree of pyramidalization at this atom. This is
consistent with electron delocalization between the N
atom and the carbonyl centre [N-C=O = 1.374(3)Å].
;
```

*Example 2 – based on C<sub>31</sub>H<sub>48</sub>N<sub>4</sub>O<sub>4</sub>, reported by Coleman, Patrick, Andersen & Rettig [Acta Cryst. (1996), C52, 1525–1527].*

```
_publ.section_title
; Hemiasterlin methyl ester
;
```

```
_publ.section_title_footnote
```

```
; IUPAC name: methyl 2,5-dimethyl-4-2-[3-methyl-
  2-methylamino-3-(N-methylbenzo[b]pyrrol-
  3-yl)butanamido]-3,3-dimethyl-N-methyl-
  butanamido-2-hexenoate.
;
```

`_publ.contact_author` (text)

`_publ_contact_author` (cif\_core.dic 2.0.1)

The name and address of the author submitting the manuscript and data block. This is the person contacted by the journal editorial staff. It is preferable to use the separate data items `_publ.contact_author_name` and `_publ.contact_author_address`.

Example:

```
; Professor George Ferguson
; Department of Chemistry and Biochemistry
; University of Guelph
; Ontario
; Canada
; N1G 2W1
```

[`publ`]

`_publ.contact_author_address` (text)

`_publ_contact_author_address` (cif\_core.dic 2.0.1)

The address of the author submitting the manuscript and data block. This is the person contacted by the journal editorial staff.

Example:

```
; Department of Chemistry and Biochemistry
; University of Guelph
; Ontario
; Canada
; N1G 2W1
```

[`publ`]

`_publ.contact_author_email` (line)

`_publ_contact_author_email` (cif\_core.dic 2.0.1)

E-mail address in a form recognizable to international networks. The format of e-mail addresses is given in Section 3.4, *Address Specification, of Internet Message Format*, RFC 2822, P. Resnick (Editor), Network Standards Group, April 2001.

Examples: 'name@host.domain.country', 'uur5@banjo.bitnet'.

[`publ`]

**\_publ.contact\_author\_fax** (line)  
\_publ\_contact\_author\_fax (cif\_core.dic 2.0.1)

Facsimile telephone number of the author submitting the manuscript and data block. The recommended style starts with the international dialing prefix, followed by the area code in parentheses, followed by the local number with no spaces. The earlier convention of including the international dialing prefix in parentheses is no longer recommended.

Examples: '12 (34) 9477330', '12 () 349477330'. [pub1]

**\_publ.contact\_author\_name** (text)  
\_publ\_contact\_author\_name (cif\_core.dic 2.0.1)

The name of the author submitting the manuscript and data block. This is the person contacted by the journal editorial staff.

Example:  
 ; Professor George Ferguson  
 ; [pub1]

**\_publ.contact\_author\_phone** (line)  
\_publ\_contact\_author\_phone (cif\_core.dic 2.0.1)

Telephone number of the author submitting the manuscript and data block. The recommended style starts with the international dialing prefix, followed by the area code in parentheses, followed by the local number and any extension number prefixed by 'x', with no spaces. The earlier convention of including the international dialing prefix in parentheses is no longer recommended.

Examples: '12 (34) 9477330', '12 () 349477330', '12 (34) 9477330x5543'. [pub1]

**\_publ.contact\_letter** (text)  
\_publ\_contact\_letter (cif\_core.dic 2.0.1)

A letter submitted to the journal editor by the contact author. [pub1]

\* **\_publ.entry\_id**  
 This data item is a pointer to \_entry.id in the ENTRY category.

**\_publ.manuscript\_creation** (text)  
\_publ\_manuscript\_creation (cif\_core.dic 2.0.1)

A description of the word-processor package and computer used to create the word-processed manuscript stored as \_publ.manuscript\_processed.

Example: "Tex file created by FrameMaker on a Sun 3/280". [pub1]

**\_publ.manuscript\_processed** (text)  
\_publ\_manuscript\_processed (cif\_core.dic 2.0.1)

The full manuscript of a paper (excluding possibly the figures and the tables) output in ASCII characters from a word processor. Information about the generation of this data item must be specified in the data item \_publ.manuscript\_creation. [pub1]

**\_publ.manuscript\_text** (text)  
\_publ\_manuscript\_text (cif\_core.dic 2.0.1)

The full manuscript of a paper (excluding figures and possibly the tables) output as standard ASCII text. [pub1]

**\_publ.requested\_category** (line)  
\_publ\_requested\_category (cif\_core.dic 2.0.1)

The category of paper submitted. For submission to *Acta Crystallographica Section C* or *Acta Crystallographica Section E*, only the codes indicated for use with these journals should be used.

The data value must be one of the following:

FA	Full article
FI	Full submission – inorganic ( <i>Acta C</i> )
FO	Full submission – organic ( <i>Acta C</i> )
FM	Full submission – metal-organic ( <i>Acta C</i> )
CI	CIF-access paper – inorganic ( <i>Acta C</i> ) (no longer in use)
CO	CIF-access paper – organic ( <i>Acta C</i> ) (no longer in use)
CM	CIF-access paper – metal-organic ( <i>Acta C</i> ) (no longer in use)
EI	Electronic submission – inorganic ( <i>Acta E</i> )
EO	Electronic submission – organic ( <i>Acta E</i> )
EM	Electronic submission – metal-organic ( <i>Acta E</i> )
AD	Addenda and Errata ( <i>Acta C</i> , <i>Acta E</i> )
SC	Short communication

Where no value is given, the assumed value is 'FA'. [pub1]

**\_publ.requested\_coeditor\_name** (line)  
\_publ\_requested\_coeditor\_name (cif\_core.dic 2.0.1)

The name of the co-editor whom the authors would like to handle the submitted manuscript. [pub1]

**\_publ.requested\_journal** (line)  
\_publ\_requested\_journal (cif\_core.dic 2.0.1)

The name of the journal to which the manuscript is being submitted. [pub1]

**\_publ.section\_abstract** (text)  
\_publ\_section\_abstract (cif\_core.dic 2.0.1)

The abstract section of a manuscript if the manuscript is submitted in parts. As an alternative see \_publ.manuscript\_text and \_publ.manuscript\_processed. [pub1]

**\_publ.section\_acknowledgements** (text)  
\_publ\_section\_acknowledgements (cif\_core.dic 2.0.1)

The acknowledgements section of a manuscript if the manuscript is submitted in parts. As an alternative see \_publ.manuscript\_text and \_publ.manuscript\_processed. [pub1]

**\_publ.section\_comment** (text)  
\_publ\_section\_comment (cif\_core.dic 2.0.1)

The comment section of a manuscript if the manuscript is submitted in parts. As an alternative see \_publ.manuscript\_text and \_publ.manuscript\_processed. [pub1]

**\_publ.section\_discussion** (text)  
\_publ\_section\_discussion (cif\_core.dic 2.0.1)

The discussion section of a manuscript if the manuscript is submitted in parts. As an alternative see \_publ.manuscript\_text and \_publ.manuscript\_processed. [pub1]

**publ.section\_experimental** (text)  
publ\_section\_experimental (cif.core.dic 2.0.1)

The experimental section of a manuscript if the manuscript is submitted in parts. As an alternative see publ.manuscript\_text and publ.manuscript\_processed. The publ.section\_exptl\_prep, publ.section\_exptl\_solution and publ.section\_exptl\_refinement items are preferred for separating the chemical preparation, structure solution and refinement aspects of the description of the experiment.

[publ]

**publ.section\_exptl\_prep** (text)  
publ\_section\_exptl\_prep (cif.core.dic 2.0.1)

The experimental preparation section of a manuscript if the manuscript is submitted in parts. As an alternative see publ.manuscript\_text and publ.manuscript\_processed.

[publ]

**publ.section\_exptl\_refinement** (text)  
publ\_section\_exptl\_refinement (cif.core.dic 2.0.1)

The experimental refinement section of a manuscript if the manuscript is submitted in parts. As an alternative see publ.manuscript\_text and publ.manuscript\_processed.

[publ]

**publ.section\_exptl\_solution** (text)  
publ\_section\_exptl\_solution (cif.core.dic 2.0.1)

The experimental solution section of a manuscript if the manuscript is submitted in parts. As an alternative see publ.manuscript\_text and publ.manuscript\_processed.

[publ]

**publ.section\_figure\_captions** (text)  
publ\_section\_figure\_captions (cif.core.dic 2.0.1)

The figure captions section of a manuscript if the manuscript is submitted in parts. As an alternative see publ.manuscript\_text and publ.manuscript\_processed.

[publ]

**publ.section\_introduction** (text)  
publ\_section\_introduction (cif.core.dic 2.0.1)

The introduction section of a manuscript if the manuscript is submitted in parts. As an alternative see publ.manuscript\_text and publ.manuscript\_processed.

[publ]

**publ.section\_references** (text)  
publ\_section\_references (cif.core.dic 2.0.1)

The references section of a manuscript if the manuscript is submitted in parts. As an alternative see publ.manuscript\_text and publ.manuscript\_processed.

[publ]

**publ.section\_synopsis** (text)  
publ\_section\_synopsis (cif.core.dic 2.0.1)

The synopsis section of a manuscript if the manuscript is submitted in parts. As an alternative see publ.manuscript\_text and publ.manuscript\_processed.

[publ]

**publ.section\_table\_legends** (text)  
publ\_section\_table\_legends (cif.core.dic 2.0.1)

The table legends section of a manuscript if the manuscript is submitted in parts. As an alternative see publ.manuscript\_text and publ.manuscript\_processed.

[publ]

**publ.section\_title** (text)  
publ\_section\_title (cif.core.dic 2.0.1)

The title of a manuscript if the manuscript is submitted in parts. As an alternative see publ.manuscript\_text and publ.manuscript\_processed.

[publ]

**publ.section\_title\_footnote** (text)  
publ\_section\_title\_footnote (cif.core.dic 2.0.1)

The footnote to the title of a manuscript if the manuscript is submitted in parts. As an alternative see publ.manuscript\_text and publ.manuscript\_processed.

[publ]

### PUBL\_AUTHOR

Data items in the PUBL\_AUTHOR category record details of the authors of a manuscript submitted for publication.

Category group(s): **inclusive\_group**

**iucr\_group**

Category key(s): **publ\_author.name**

*Example 1 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].*

```
loop_
  publ_author.name
  publ_author.address
  'Willis, Anthony C.'
;
  Research School of Chemistry
  Australian National University
  GPO Box 4
  Canberra, A.C.T.
  Australia      2601
;
```

**publ\_author.address** (text)  
publ\_author\_address (cif.core.dic 2.0.1)

The address of a publication author. If there is more than one author this is looped with publ\_author.name.

Example:

```
; Department
  Institute
  Street
  City and postcode
  COUNTRY
```

[publ\_author]

**publ\_author.email** (text)  
publ\_author\_email (cif.core.dic 2.3.1)

The e-mail address of a publication author. If there is more than one author, this will be looped with publ\_author.name. The format of e-mail addresses is given in Section 3.4, *Address Specification*, of *Internet Message Format*, RFC 2822, P. Resnick (Editor), Network Standards Group, April 2001.

Examples: 'name@host.domain.country', 'bm@iucr.org'. [publ\_author]

**publ\_author.footnote** (line)  
publ\_author\_footnote (cif.core.dic 2.0.1)

A footnote accompanying an author's name in the list of authors of a paper. Typically indicates sabbatical address, additional affiliations or date of decease.

Examples: 'On leave from U. Western Australia',  
 'Also at Department of Biophysics'.

[publ\_author]

**publ\_author.id\_iucr** (code)  
publ\_author\_id\_iucr (cif.core.dic 2.3)  
 Identifier in the IUCr contact database of a publication author. This identifier may be available from the *World Directory of Crystallographers* (<http://wdc.iucr.org>).  
 Example: '2985'. [publ\_author]

\* **publ\_author.name** (line)  
publ\_author\_name (cif.core.dic 2.0.1)  
 The name of a publication author. If there are multiple authors this will be looped with publ\_author.address. The family name(s), followed by a comma and including any dynastic components, precedes the first names or initials.  
 Examples: 'Bleary, Percival R.', 'O'Neil, F.K.', 'Van den Bossche, G.', 'Yang, D.-L.', 'Simonov, Yu.A.'. [publ\_author]

**publ\_body.contents** (text)  
publ\_body\_contents (cif.core.dic 2.0.1)  
 A text section of a paper. [publ\_body]

**publ\_body.element** (code)  
publ\_body\_element (cif.core.dic 2.0.1)  
 The functional role of the associated text section.  
 The data value must be one of the following:  
 section  
 subsection  
 subsubsection  
 appendix  
 footnote [publ\_body]

**publ\_body.format** (code)  
publ\_body\_format (cif.core.dic 2.0.1)  
 Code indicating the appropriate typesetting conventions for accented characters and special symbols in the text section.  
 The data value must be one of the following:  
 ascii no coding for special symbols  
 cif CIF convention  
 latex LaTeX  
 sgml SGML (ISO 8879)  
 tex TeX  
 troff troff or nroff [publ\_body]

**PUBL\_BODY**

Data items in the PUBL\_BODY category permit the labelling of different text sections within the body of a paper. Note that these should not be used in a paper which has a standard format with sections tagged by specific data names (such as in *Acta Crystallographica Section C*). Typically, each journal will supply a list of the specific items it requires in its *Notes for Authors*.  
 Category group(s): inclusive\_group  
                   iucr\_group  
 Category key(s): publ\_body.element  
                   publ\_body.label

---

*Example 1 – based on a paper by R. Restori & D. Schwarzenbach [Acta Cryst. (1996), A52, 369–378].*

```
loop_
  _publ_body.element
  _publ_body.label
  _publ_body.title
  _publ_body.format
  _publ_body.contents

  section 1 Introduction cif
; X-ray diffraction from a crystalline material provides
  information on the thermally and spatially averaged
  electron density in the crystal...
;
  section 2 Theory tex
; In the rigid-atom approximation, the dynamic electron
  density of an atom is described by the convolution
  product of the static atomic density and a probability
  density function,
  $\rho_{\text{dyn}}(\mathbf{r}) = \rho_{\text{stat}}(\mathbf{r}) * P(\mathbf{r})$. \eqno(1)$
;
Example 2 – based on a paper by R. J. Papoular, Y. Vekhter & P. Coppens [Acta Cryst. (1996), A52, 397–407].
```

```
loop_
  _publ_body.element
  _publ_body.label
  _publ_body.title
  _publ_body.contents

  section 3
; The two-channel method for retrieval of the deformation
  electron density
;
  subsection 3.1 'The two-channel entropy S[D\r(r)]'
; As the wide dynamic range involved in the total electron
  density...
;
  subsection 3.2
'Uniform vs informative prior model densities'
  subsubsection 3.2.1 'Use of uniform models'
; Straightforward algebra leads to expressions analogous
  to...
;
```

**publ\_body.label** (code)  
publ\_body\_label (cif.core.dic 2.0.1)  
 Code identifying the section of text.  
 Examples: '1', '1.1', '2.1.3'. [publ\_body]

**publ\_body.title** (text)  
publ\_body\_title (cif.core.dic 2.0.1)  
 Title of the associated section of text. [publ\_body]

**PUBL\_MANUSCRIPT\_INCL**

Data items in the PUBL\_MANUSCRIPT\_INCL category allow the authors of a manuscript submitted for publication to list data names that should be added to the standard request list used by the journal printing software.  
 Category group(s): inclusive\_group  
                   iucr\_group  
 Category key(s): publ\_manuscript\_incl.entry\_id

---

*Example 1 – hypothetical example.*

```
_publ_manuscript_incl.entry_id 'EXAMHYPO'
loop_
  _publ_manuscript_incl.extra_item
  _publ_manuscript_incl.extra_info
  _publ_manuscript_incl.extra_defn
'atom_site.symmetry_multiplicity'
  'to emphasise special sites' yes
'_chemical.compound_source'
  'rare material, unusual source' yes
'_reflns.d_resolution_high'
  'limited data are a problem here' yes
'_crystal.magnetic_permeability'
  'unusual value for this material' no
```

\* **publ\_manuscript\_incl.entry\_id**  
 This data item is a pointer to entry\_id in the ENTRY category.

**publ\_manuscript\_incl.extra\_defn** (line)

publ\_manuscript\_incl\_extra\_defn(cif\_core.dic 2.0.1)

Flags whether the corresponding data item marked for inclusion in a journal request list is a standard CIF definition or not.

The data value must be one of the following:

no	not a standard CIF data name
n	abbreviation for 'no'
yes	a standard CIF data name
y	abbreviation for 'yes'

[publ\_manuscript\_incl]

**publ\_manuscript\_incl.extra\_info** (text)

publ\_manuscript\_incl\_extra\_info(cif\_core.dic 2.0.1)

A short note indicating the reason why the author wishes the corresponding data item marked for inclusion in the journal request list to be published.

Examples: 'to emphasise very special sites',

'rare material from unusual source',

'the limited data is a problem here',

'a new data quantity needed here'.

[publ\_manuscript\_incl]

**publ\_manuscript\_incl.extra\_item** (line)

publ\_manuscript\_incl\_extra\_item(cif\_core.dic 2.0.1)

Specifies the inclusion of specific data into a manuscript which are not normally requested by the journal. The values of this item are the extra data names (which *must* be enclosed in single quotes) that will be added to the journal request list.

Examples: 'atom\_site.symmetry\_multiplicity',

'chemical.compound.source', 'reflns.d.resolution.high',

'crystal.magnetic.permeability'.

[publ\_manuscript\_incl]

## REFINE

Data items in the REFINE category record details about the structure-refinement parameters.

Category group(s): **inclusive\_group**

**refine\_group**

Category key(s): **refine.entry\_id**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```

_refine.entry_id          '5HVP'
_refine.ls_number_reflns_obs 12901
_refine.ls_number_restraints 6609
_refine.ls_number_parameters 7032
_refine.ls_R_Factor_obs     0.176
_refine.ls_weighting_scheme  calc
_refine.ls_weighting_details
; Sigdel model of Konnert-Hendrickson:
  Sigdel: Afsig + Bfsig*(sin(theta)/lambda-1/6)
  Afsig = 22.0, Bfsig = -150.0 at beginning of refinement
  Afsig = 15.5, Bfsig = -50.0 at end of refinement
;

```

*Example 2 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].*

```

_refine.details          sfls: F_calc_weight_full_matrix

_refine.ls_structure_factor_coef F
_refine.ls_matrix_type    full
_refine.ls_weighting_scheme  calc
_refine.ls_weighting_details 'w=1/(\s^2(F)+0.0004F^2)'
_refine.ls_hydrogen_treatment 'refxyz except H332B noref'
_refine.ls_extinction_method Zachariasen
_refine.ls_extinction_coef  3514
_refine.ls_extinction_expression
; Larson, A. C. (1970). "Crystallographic Computing", edited
  by F. R. Ahmed. Eq. (22) p. 292. Copenhagen: Munksgaard.

_refine.ls_abs_structure_details
; The absolute configuration was assigned to agree with the
  known chirality at C3 arising from its precursor l-leucine.
;
_refine.ls_abs_structure_Flack      0
_refine.ls_number_reflns_obs        1408
_refine.ls_number_parameters        272
_refine.ls_number_restraints        0
_refine.ls_number_constraints       0
_refine.ls_R_factor_all              .038
_refine.ls_R_factor_obs              .034
_refine.ls_wR_factor_all             .044
_refine.ls_wR_factor_obs             .042
_refine.ls_goodness_of_fit_all       1.462
_refine.ls_goodness_of_fit_obs       1.515
_refine.ls_shift_over_esd_max        .535
_refine.ls_shift_over_esd_mean       .044
_refine.diff_density_min             -.108
_refine.diff_density_max             .131

```

**refine.aniso\_B[1][1]** (float)

The [1][1] element of the matrix that defines the overall anisotropic displacement model if one was refined for this structure.

[refine]

**refine.aniso\_B[1][2]** (float)

The [1][2] element of the matrix that defines the overall anisotropic displacement model if one was refined for this structure.

[refine]

**refine.aniso\_B[1][3]** (float)

The [1][3] element of the matrix that defines the overall anisotropic displacement model if one was refined for this structure.

[refine]

**refine.aniso\_B[2][2]** (float)

The [2][2] element of the matrix that defines the overall anisotropic displacement model if one was refined for this structure.

[refine]

**refine.aniso\_B[2][3]** (float)

The [2][3] element of the matrix that defines the overall anisotropic displacement model if one was refined for this structure.

[refine]

**refine.aniso\_B[3][3]** (float)

The [3][3] element of the matrix that defines the overall anisotropic displacement model if one was refined for this structure.

[refine]

**refine.B\_iso\_max** (float)

The maximum isotropic displacement parameter (*B* value) found in the coordinate set.

[refine]

**`_refine.B_iso_mean`** (float)  
The mean isotropic displacement parameter (*B* value) for the coordinate set.

[refine]

**`_refine.B_iso_min`** (float)  
The minimum isotropic displacement parameter (*B* value) found in the coordinate set.

[refine]

**`_refine.correlation_coeff_Fo_to_Fc`** (float)  
**`_refine.ebi_Correlation_coeff_Fo_to_Fc`** (*ebi-extensions 1.0*)  
The correlation coefficient between the observed and calculated structure factors for reflections included in the refinement. The correlation coefficient is scale-independent and gives an idea of the quality of the refined model.

$$R_{\text{corr}} = \frac{\sum_i (F_{oi} F_{ci} - \langle F_o \rangle \langle F_c \rangle)}{\sqrt{\sum_i (F_{oi})^2 - \langle F_o \rangle^2} \sqrt{\sum_i (F_{ci})^2 - \langle F_c \rangle^2}},$$

where  $F_o$  = observed structure factors,  $F_c$  = calculated structure factors,  $\langle \rangle$  denotes average value and the summation is over reflections included in the refinement.

[refine]

**`_refine.correlation_coeff_Fo_to_Fc_free`** (float)  
**`_refine.ebi_Correlation_coeff_Fo_to_Fc_free`** (*ebi-extensions 1.0*)  
The correlation coefficient between the observed and calculated structure factors for reflections not included in the refinement (free reflections). The correlation coefficient is scale-independent and gives an idea of the quality of the refined model.

$$R_{\text{corr}} = \frac{\sum_i (F_{oi} F_{ci} - \langle F_o \rangle \langle F_c \rangle)}{\sqrt{\sum_i (F_{oi})^2 - \langle F_o \rangle^2} \sqrt{\sum_i (F_{ci})^2 - \langle F_c \rangle^2}},$$

where  $F_o$  = observed structure factors,  $F_c$  = calculated structure factors,  $\langle \rangle$  denotes average value and the summation is over reflections not included in the refinement (free reflections).

[refine]

**`_refine.details`** (text)  
**`_refine.special_details`** (*cif-core.dic 2.0.1*)  
Description of special aspects of the refinement process.

[refine]

**`_refine.diff_density_max`** (float, su)  
**`_refine.diff_density_max`** (*cif-core.dic 2.0.1*)  
The maximum value of the electron density in the final difference Fourier map.

Related item: **`_refine.diff_density_max_esd`** (associated esd). [refine]

**`_refine.diff_density_max_esd`** (float)  
The standard uncertainty (estimated standard deviation) of **`_refine.diff_density_max`**.

Related item: **`_refine.diff_density_max`** (associated value). [refine]

**`_refine.diff_density_min`** (float, su)  
**`_refine.diff_density_min`** (*cif-core.dic 2.0.1*)  
The minimum value of the electron density in the final difference Fourier map.

Related item: **`_refine.diff_density_min_esd`** (associated esd). [refine]

**`_refine.diff_density_min_esd`** (float)  
The standard uncertainty (estimated standard deviation) of **`_refine.diff_density_min`**.

Related item: **`_refine.diff_density_min`** (associated value). [refine]

**`_refine.diff_density_rms`** (float, su)  
**`_refine.diff_density_rms`** (*cif-core.dic 2.0.1*)  
The root-mean-square-deviation of the electron density in the final difference Fourier map. This value is measured with respect to the arithmetic mean density and is derived from summations over each grid point in the asymmetric unit of the cell. This quantity is useful for assessing the significance of the values of **`_refine.diff_density_min`** and **`_refine.diff_density_max`**, and also for defining suitable contour levels.

Related item: **`_refine.diff_density_rms_esd`** (associated esd). [refine]

**`_refine.diff_density_rms_esd`** (float)  
The standard uncertainty (estimated standard deviation) of **`_refine.diff_density_rms`**.

Related item: **`_refine.diff_density_rms`** (associated value). [refine]

\* **`_refine.entry_id`**  
This data item is a pointer to **`_entry.id`** in the ENTRY category.

**`_refine.ls_abs_structure_details`** (text)  
**`_refine.ls_abs_structure_details`** (*cif-core.dic 2.0.1*)  
The nature of the absolute structure and how it was determined. For example, this may describe the Friedel pairs used.

[refine]

**`_refine.ls_abs_structure_Flack`** (float, su)  
**`_refine.ls_abs_structure_Flack`** (*cif-core.dic 2.0.1*)

The measure of absolute structure (enantiomorph or polarity) as defined by Flack (1983). For centrosymmetric structures the only permitted value, if the data name is present, is 'inapplicable', represented by '.'. For noncentrosymmetric structures the value must lie in the 99.97% Gaussian confidence interval  $-3u \leq x \leq 1 + 3u$  and a standard uncertainty (estimated standard deviation)  $u$  must be supplied. The item range of [0.0, 1.0] is correctly interpreted as meaning  $(0.0 - 3u) \leq x \leq (1.0 + 3u)$ .

Reference: Flack, H. D. (1983). *Acta Cryst.* **A39**, 876–881.

The permitted range is [0.0, 1.0].

Related item: **`_refine.ls_abs_structure_Flack_esd`** (associated esd). [refine]

**`_refine.ls_abs_structure_Flack_esd`** (float)  
The standard uncertainty (estimated standard deviation) of **`_refine.ls_abs_structure_Flack`**.

Related item: **`_refine.ls_abs_structure_Flack`** (associated value). [refine]

**`_refine.ls_abs_structure_Rogers`** (float, su)  
**`_refine.ls_abs_structure_Rogers`** (*cif-core.dic 2.0.1*)

The measure of absolute structure (enantiomorph or polarity) as defined by Rogers. The value must lie in the 99.97% Gaussian confidence interval  $-1 - 3u \leq \eta \leq 1 + 3u$  and a standard uncertainty (estimated standard deviation)  $u$  must be supplied. The item range of [-1.0, 1.0] is correctly interpreted as meaning  $(-1.0 - 3u) \leq \eta \leq (1.0 + 3u)$ .

Reference: Rogers, D. (1981). *Acta Cryst.* **A37**, 734–741.

The permitted range is [-1.0, 1.0].

Related item: **`_refine.ls_abs_structure_Rogers_esd`** (associated esd). [refine]

**`_refine.ls_abs_structure_Rogers_esd`** (float)  
The standard uncertainty (estimated standard deviation) of `_refine.ls_abs_structure_Rogers`.  
Related item: `_refine.ls_abs_structure_Rogers` (associated value). [refine]

\* **`_refine.ls_d_res_high`** (float)  
`_refine_ls_d_res_high` (cif\_core.dic 2.0.1)  
The smallest value for the interplanar spacings for the reflection data used in the refinement in ångströms. This is called the highest resolution.  
The permitted range is [0.0, ∞). [refine]

\* **`_refine.ls_d_res_low`** (float)  
`_refine_ls_d_res_low` (cif\_core.dic 2.0.1)  
The largest value for the interplanar spacings for the reflection data used in the refinement in ångströms. This is called the lowest resolution.  
The permitted range is [0.0, ∞). [refine]

**`_refine.ls_extinction_coef`** (float, su)  
`_refine_ls_extinction_coef` (cif\_core.dic 2.0.1)  
The extinction coefficient used to calculate the correction factor applied to the structure-factor data. The nature of the extinction coefficient is given in the definitions of `_refine.ls_extinction_expression` and `_refine.ls_extinction_method`. For the 'Zachariasen' method it is the  $r^*$  value; for the 'Becker-Coppens type 1 isotropic' method it is the 'g' value, and for 'Becker-Coppens type 2 isotropic' corrections it is the 'ρ' value. Note that the magnitude of these values is usually of the order of 10 000.  
References: Becker, P. J. & Coppens, P. (1974). *Acta Cryst.* **A30**, 129–147, 148–153. Zachariasen, W. H. (1967). *Acta Cryst.* **23**, 558–564. Larson, A. C. (1967). *Acta Cryst.* **23**, 664–665.  
Related item: `_refine.ls_extinction_coef_esd` (associated esd).  
Example: '3472' (Zachariasen coefficient  $r^* = 0.347\text{ E}04$ ). [refine]

**`_refine.ls_extinction_coef_esd`** (float)  
The standard uncertainty (estimated standard deviation) of `_refine.ls_extinction_coef`.  
Related item: `_refine.ls_extinction_coef` (associated value). [refine]

**`_refine.ls_extinction_expression`** (text)  
`_refine_ls_extinction_expression` (cif\_core.dic 2.0.1)  
A description of or reference to the extinction-correction equation used to apply the data item `_refine.ls_extinction_coef`. This information must be sufficient to reproduce the extinction-correction factors applied to the structure factors.  
Example:  
; Larson, A. C. (1970). "Crystallographic Computing", edited by F. R. Ahmed. Eq. (22), p.292. Copenhagen: Munksgaard.  
; [refine]

**`_refine.ls_extinction_method`** (text)  
`_refine_ls_extinction_method` (cif\_core.dic 2.0.1)  
A description of the extinction-correction method applied. This description should include information about the correction method, either 'Becker-Coppens' or 'Zachariasen'. The latter is sometimes referred to as the 'Larson' method even though it employs Zachariasen's formula. The Becker-Coppens procedure is referred to as 'type 1' when correcting secondary extinction dominated by the mosaic spread; as 'type 2' when secondary extinction is dominated by particle size and includes a primary extinction component; and as 'mixed' when there is a mixture of types 1 and 2. For the Becker-Coppens method, it is also necessary to set the mosaic distribution as either 'Gaussian' or 'Lorentzian' and the nature of the extinction as 'isotropic' or

'anisotropic'. Note that if either the 'mixed' or 'anisotropic' corrections are applied the multiple coefficients cannot be contained in `*_extinction_coef` and must be listed in `_refine.details`.

References: Becker, P. J. & Coppens, P. (1974). *Acta Cryst.* **A30**, 129–147, 148–153. Zachariasen, W. H. (1967). *Acta Cryst.* **23**, 558–564. Larson, A. C. (1967). *Acta Cryst.* **23**, 664–665.  
Example: 'B-C type 2 Gaussian isotropic'. [refine]

**`_refine.ls_goodness_of_fit_all`** (float, su)  
`_refine_ls_goodness_of_fit_all` (cif\_core.dic 2.0.1)  
The least-squares goodness-of-fit parameter  $S$  for all data after the final cycle of refinement. Ideally, account should be taken of parameters restrained in the least-squares refinement. See also the definition of `_refine.ls_restrained_s_all`.

$$S = \left( \frac{\sum |w| |Y_{\text{obs}} - Y_{\text{calc}}|^2}{N_{\text{ref}} - N_{\text{param}}} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed coefficients (see `_refine.ls_structure_factor_coef`),  $Y_{\text{calc}}$  = the calculated coefficients (see `_refine.ls_structure_factor_coef`),  $w$  = the least-squares reflection weight [ $1/(\text{e.s.d. squared})$ ],  $N_{\text{ref}}$  = the number of reflections used in the refinement and  $N_{\text{param}}$  = the number of refined parameters; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_goodness_of_fit_all_esd` (associated esd). [refine]

**`_refine.ls_goodness_of_fit_all_esd`** (float)  
The standard uncertainty (estimated standard deviation) of `_refine.ls_goodness_of_fit_all`.  
Related item: `_refine.ls_goodness_of_fit_all` (associated value). [refine]

**`_refine.ls_goodness_of_fit_gt`** (float)  
`_refine_ls_goodness_of_fit_gt` (cif\_core.dic 2.3)  
The least-squares goodness-of-fit parameter  $S$  for significantly intense reflections (see `_reflns.threshold_expression`) after the final cycle of refinement. Ideally, account should be taken of parameters restrained in the least-squares refinement. See also `_refine.ls_restrained_s_definitions`.

$$S = \left( \frac{\sum |w| |Y_{\text{obs}} - Y_{\text{calc}}|^2}{N_{\text{ref}} - N_{\text{param}}} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed coefficients (see `_refine.ls_structure_factor_coef`),  $Y_{\text{calc}}$  = the calculated coefficients (see `_refine.ls_structure_factor_coef`),  $w$  = the least-squares reflection weight [ $1/(u^2)$ ],  $u$  = the standard uncertainty,  $N_{\text{ref}}$  = the number of reflections used in the refinement and  $N_{\text{param}}$  = the number of refined parameters; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_goodness_of_fit_obs` (alternate). [refine]

**`_refine.ls_goodness_of_fit_obs`** (float, su)  
`_refine_ls_goodness_of_fit_obs` (cif\_core.dic 2.0.1)

The least-squares goodness-of-fit parameter  $S$  for reflection data classified as ‘observed’ (see `_reflns.observed_criterion`) after the final cycle of refinement. Ideally, account should be taken of parameters restrained in the least-squares refinement. See also the definition of `_refine.ls_restrained_s_obs`.

$$S = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2|}{N_{\text{ref}} - N_{\text{param}}} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed coefficients (see `_refine.ls_structure_factor_coef`),  $Y_{\text{calc}}$  = the calculated coefficients (see `_refine.ls_structure_factor_coef`),  $w$  = the least-squares reflection weight [ $1/(\text{e.s.d. squared})$ ],  $N_{\text{ref}}$  = the number of reflections used in the refinement and  $N_{\text{param}}$  = the number of refined parameters; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_goodness_of_fit_obs_esd` (associated esd). [refine]

**`_refine.ls_goodness_of_fit_obs_esd`** (float)

The standard uncertainty (estimated standard deviation) of `_refine.ls_goodness_of_fit_obs`.

Related item: `_refine.ls_goodness_of_fit_obs` (associated value). [refine]

**`_refine.ls_goodness_of_fit_ref`** (float)

`_refine_ls_goodness_of_fit_ref` (cif\_core.dic 2.3)

The least-squares goodness-of-fit parameter  $S$  for all reflections included in the refinement after the final cycle of refinement. Ideally, account should be taken of parameters restrained in the least-squares refinement. See also `_refine.ls_restrained_s` definitions.

$$S = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2|}{N_{\text{ref}} - N_{\text{param}}} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed coefficients (see `_refine.ls_structure_factor_coef`),  $Y_{\text{calc}}$  = the calculated coefficients (see `_refine.ls_structure_factor_coef`),  $w$  = the least-squares reflection weight [ $1/(u^2)$ ],  $u$  = the standard uncertainty,  $N_{\text{ref}}$  = the number of reflections used in the refinement and  $N_{\text{param}}$  = the number of refined parameters; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞). [refine]

**`_refine.ls_hydrogen_treatment`** (ucode)

`_refine_ls_hydrogen_treatment` (cif\_core.dic 2.0.1)

Treatment of hydrogen atoms in the least-squares refinement.

The data value must be one of the following:

refall refined all H-atom parameters  
 refxyz refined H-atom coordinates only  
 refU refined H-atom  $U$ 's only  
 noref no refinement of H-atom parameters  
 constr H-atom parameters constrained  
 mixed some constrained, some independent  
 undef H-atom parameters not defined

[refine]

**`_refine.ls_matrix_type`** (ucode)

`_refine_ls_matrix_type` (cif\_core.dic 2.0.1)

Type of matrix used to accumulate the least-squares derivatives.

The data value must be one of the following:

full full  
 fullcycle full with fixed elements per cycle

atomblock block diagonal per atom  
 userblock user-defined blocks  
 diagonal diagonal elements only  
 sparse selected elements only

[refine]

**`_refine.ls_number_constraints`** (int)

`_refine_ls_number_constraints` (cif\_core.dic 2.0.1)

The number of constrained (non-refined or dependent) parameters in the least-squares process. These may be due to symmetry or any other constraint process (e.g. rigid-body refinement). See also `_atom_site.constraints` and `_atom_site.refinement_flags`. A general description of constraints may appear in `_refine.details`.

The permitted range is [0, ∞). [refine]

**`_refine.ls_number_parameters`** (int)

`_refine_ls_number_parameters` (cif\_core.dic 2.0.1)

The number of parameters refined in the least-squares process. If possible, this number should include some contribution from the restrained parameters. The restrained parameters are distinct from the constrained parameters (where one or more parameters are linearly dependent on the refined value of another). Least-squares restraints often depend on geometry or energy considerations and this makes their direct contribution to this number, and to the goodness-of-fit calculation, difficult to assess.

The permitted range is [0, ∞). [refine]

**`_refine.ls_number_reflns_all`** (int)

The number of reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low`.

The permitted range is [0, ∞). [refine]

**`_refine.ls_number_reflns_obs`** (int)

`_refine_ls_number_reflns` (cif\_core.dic 2.0.1)

The number of reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflns.observed_criterion`.

The permitted range is [0, ∞). [refine]

**`_refine.ls_number_reflns_R_free`** (int)

The number of reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the test reflections (i.e. were excluded from the refinement) when the refinement included the calculation of a ‘free’  $R$  factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`.

The permitted range is [0, ∞). [refine]

**`_refine.ls_number_reflns_R_work`** (int)

The number of reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the working reflections (i.e. were included in the refinement) when the refinement included the calculation of a ‘free’  $R$  factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`.

The permitted range is [0, ∞). [refine]

**`_refine.ls_number_restraints`** (int)  
`_refine_ls_number_restraints` (cif\_core.dic 2.0.1)

The number of restrained parameters. These are parameters which are not directly dependent on another refined parameter. Restrained parameters often involve geometry or energy dependencies. See also `_atom_site.constraints` and `_atom_site.refinement_flags`. A general description of refinement constraints may appear in `_refine.details`.

The permitted range is [0, ∞). [refine]

**`_refine.ls_percent_reflms_obs`** (float)

The number of reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflms.observed_criterion`, expressed as a percentage of the number of geometrically observable reflections that satisfy the resolution limits.

[refine]

**`_refine.ls_percent_reflms_R_free`** (float)

The number of reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflms.observed_criterion`, and that were used as the test reflections (*i.e.* were excluded from the refinement) when the refinement included the calculation of a ‘free’ *R* factor, expressed as a percentage of the number of geometrically observable reflections that satisfy the resolution limits.

[refine]

**`_refine.ls_R_factor_all`** (float)

`_refine_ls_R_factor_all` (cif\_core.dic 2.0.1)

Residual factor *R* for all reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|},$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_wR_factor_all` (alternate). [refine]

**`_refine.ls_R_factor_gt`** (float)

`_refine_ls_R_factor_gt` (cif\_core.dic 2.3)

Residual factor for the reflections (with number given by `_reflms.number_gt`) judged significantly intense (*i.e.* satisfying the threshold specified by `_reflms.threshold_expression`) and included in the refinement. The reflections also satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low`. This is the conventional *R* factor. See also `_refine.ls_wR_factor_definitions`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|},$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_R_factor_obs` (alternate). [refine]

**`_refine.ls_R_factor_obs`** (float)

`_refine_ls_R_factor_obs` (cif\_core.dic 2.0.1)

Residual factor *R* for reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflms.observed_criterion`. `_refine.ls_R_factor_obs` should not be confused with `_refine.ls_R_factor_R_work`; the former reports the results of a refinement in which all observed reflections were used, the latter a refinement in which a subset of the observed reflections were excluded from refinement for the calculation of a ‘free’ *R* factor. However, it would be meaningful to quote both values if a ‘free’ *R* factor were calculated for most of the refinement, but all of the observed reflections were used in the final rounds of refinement; such a protocol should be explained in `_refine.details`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|},$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_wR_factor_obs` (alternate). [refine]

**`_refine.ls_R_factor_R_free`** (float)

Residual factor *R* for reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflms.observed_criterion`, and that were used as the test reflections (*i.e.* were excluded from the refinement) when the refinement included the calculation of a ‘free’ *R* factor. Details of how reflections were assigned to the working and test sets are given in `_reflms.R_free_details`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|},$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related items: `_refine.ls_wR_factor_R_free` (alternate), `_refine.ls_R_factor_R_free_error` (associated error). [refine]

**`_refine.ls_R_factor_R_free_error`** (float)

The estimated error in `_refine.ls_R_factor_R_free`. The method used to estimate the error is described in the item `_refine.ls_R_factor_R_free_error_details`.

Related item: `_refine.ls_R_factor_R_free` (associated value). [refine]

**`_refine.ls_R_factor_R_free_error_details`** (text)

Special aspects of the method used to estimate the error in `_refine.ls_R_factor_R_free`.

[refine]

**`_refine.ls_R_factor_R_work`** (float)  
 Residual factor  $R$  for reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the working reflections (*i.e.* were included in the refinement) when the refinement included the calculation of a ‘free’  $R$  factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`. `_refine.ls_R_factor_obs` should not be confused with `_refine.ls_R_factor_R_work`; the former reports the results of a refinement in which all observed reflections were used, the latter a refinement in which a subset of the observed reflections were excluded from refinement for the calculation of a ‘free’  $R$  factor. However, it would be meaningful to quote both values if a ‘free’  $R$  factor were calculated for most of the refinement, but all of the observed reflections were used in the final rounds of refinement; such a protocol should be explained in `_refine.details`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|},$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_wR_factor_R_work` (alternate). [refine]

**`_refine.ls_R_Fsqd_factor_obs`** (float)  
`_refine_ls_R_Fsqd_factor` (cif-core.dic 2.0.1)  
 Residual factor  $R(F^2)$  for reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflns.observed_criterion`, calculated on the squares of the observed and calculated structure-factor amplitudes.

$$R(F^2) = \frac{\sum |F_{\text{obs}}^2 - F_{\text{calc}}^2|}{\sum |F_{\text{obs}}^2|},$$

where  $F_{\text{obs}}^2$  = squares of the observed structure-factor amplitudes,  $F_{\text{calc}}^2$  = squares of the calculated structure-factor amplitudes and the sum is taken over the specified reflections.

The permitted range is [0.0, ∞). [refine]

**`_refine.ls_R_I_factor_obs`** (float)  
`_refine_ls_R_I_factor` (cif-core.dic 2.0.1)  
 Residual factor  $R(I)$  for reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflns.observed_criterion`, calculated on the estimated reflection intensities. This is most often calculated in Rietveld refinements against powder data, where it is referred to as  $R_B$  or  $R_{\text{Bragg}}$ .

$$R(I) = \frac{\sum |I_{\text{obs}} - I_{\text{calc}}|}{\sum |I_{\text{obs}}|},$$

where  $I_{\text{obs}}$  = the net observed intensities,  $I_{\text{calc}}$  = the net calculated intensities and the sum is taken over the specified reflections.

The permitted range is [0.0, ∞). [refine]

**`_refine.ls_redundancy_reflms_all`** (float)  
 The ratio of the total number of observations of the reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` to the number of crystallographically unique reflections that satisfy the same limits.

[refine]

**`_refine.ls_redundancy_reflms_obs`** (float)  
 The ratio of the total number of observations of the reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflns.observed_criterion` to the number of crystallographically unique reflections that satisfy the same limits.

[refine]

**`_refine.ls_restrained_S_all`** (float)  
`_refine_ls_restrained_S_all` (cif-core.dic 2.0.1)  
 The least-squares goodness-of-fit parameter  $S'$  for all reflections after the final cycle of least-squares refinement. This parameter explicitly includes the restraints applied in the least-squares process. See also the definition of `_refine.ls_goodness_of_fit_all`.

$$S' = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2 + \sum_r |w_r|P_{\text{calc}} - P_{\text{targ}}|^2}{N_{\text{ref}} + N_{\text{restr}} - N_{\text{param}}} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed coefficients (see `_refine.ls_structure_factor_coef`),  $Y_{\text{calc}}$  = the calculated coefficients (see `_refine.ls_structure_factor_coef`),  $w$  = the least-squares reflection weight [1/(e.s.d. squared)],  $P_{\text{calc}}$  = the calculated restraint values,  $P_{\text{targ}}$  = the target restraint values,  $w_r$  = the restraint weight,  $N_{\text{ref}}$  = the number of reflections used in the refinement (see `_refine.ls_number_reflms_obs`),  $N_{\text{restr}}$  = the number of restraints (see `_refine.ls_number_restraints`) and  $N_{\text{param}}$  = the number of refined parameters (see `_refine.ls_number_parameters`); the sum  $\sum$  is taken over the specified reflections and the sum  $\sum_r$  is taken over the restraints.

The permitted range is [0.0, ∞). [refine]

**`_refine.ls_restrained_S_obs`** (float)  
`_refine_ls_restrained_S_obs` (cif-core.dic 2.0.1)

The least-squares goodness-of-fit parameter  $S'$  for reflection data classified as observed (see `_reflns.observed_criterion`) after the final cycle of least-squares refinement. This parameter explicitly includes the restraints applied in the least-squares process. See also the definition of `_refine.ls_goodness_of_fit_obs`.

$$S' = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2 + \sum_r |w_r|P_{\text{calc}} - P_{\text{targ}}|^2}{N_{\text{ref}} + N_{\text{restr}} - N_{\text{param}}} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed coefficients (see `_refine.ls_structure_factor_coef`),  $Y_{\text{calc}}$  = the calculated coefficients (see `_refine.ls_structure_factor_coef`),  $w$  = the least-squares reflection weight [1/(e.s.d. squared)],  $P_{\text{calc}}$  = the calculated restraint values,  $P_{\text{targ}}$  = the target restraint values,  $w_r$  = the restraint weight,  $N_{\text{ref}}$  = the number of reflections used in the refinement (see `_refine.ls_number_reflms_obs`),  $N_{\text{restr}}$  = the number of restraints (see `_refine.ls_number_restraints`) and  $N_{\text{param}}$  = the number of refined parameters (see `_refine.ls_number_parameters`); the sum  $\sum$  is taken over the specified reflections and the sum  $\sum_r$  is taken over the restraints.

The permitted range is [0.0, ∞). [refine]

**`_refine.ls_shift_over_esd_max`** (float)  
`_refine_ls_shift/esd_max` (cif-core.dic 2.0.1)

The largest ratio of the final least-squares parameter shift to the final standard uncertainty (estimated standard deviation).

The permitted range is [0.0, ∞). [refine]

**`_refine.ls_shift_over_esd_mean`** (float)  
`_refine_ls_shift/esd_mean` (cif\_core.dic 2.0.1)

The average ratio of the final least-squares parameter shift to the final standard uncertainty (estimated standard deviation).

The permitted range is [0.0, ∞). [refine]

**`_refine.ls_shift_over_su_max`** (float)  
`_refine_ls_shift/su_max` (cif\_core.dic 2.3)

The largest ratio of the final least-squares parameter shift to the final standard uncertainty. [refine]

The permitted range is [0.0, ∞).

Related item: `_refine.ls_shift_over_esd_max` (alternate). [refine]

**`_refine.ls_shift_over_su_max_lt`** (float)  
`_refine_ls_shift/su_max_lt` (cif\_core.dic 2.3)

An upper limit for the largest ratio of the final least-squares parameter shift to the final standard uncertainty. This item is used when the largest value of the shift divided by the final standard uncertainty is too small to measure.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_shift_over_su_max` (alternate). [refine]

**`_refine.ls_shift_over_su_mean`** (float)  
`_refine_ls_shift/su_mean` (cif\_core.dic 2.3)

The average ratio of the final least-squares parameter shift to the final standard uncertainty.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_shift_over_esd_mean` (alternate). [refine]

**`_refine.ls_shift_over_su_mean_lt`** (float)  
`_refine_ls_shift/su_mean_lt` (cif\_core.dic 2.3)

An upper limit for the average ratio of the final least-squares parameter shift to the final standard uncertainty. This item is used when the average value of the shift divided by the final standard uncertainty is too small to measure.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_shift_over_su_mean` (alternate). [refine]

**`_refine.ls_structure_factor_coef`** (ucode)  
`_refine_ls_structure_factor_coef` (cif\_core.dic 2.0.1)

Structure-factor coefficient  $|F|$ ,  $F^2$  or  $I$  used in the least-squares refinement process.

The data value must be one of the following:

F	structure-factor magnitude
Fsqd	structure factor squared
Inet	net intensity

[refine]

**`_refine.ls_weighting_details`** (text)  
`_refine_ls_weighting_details` (cif\_core.dic 2.0.1)

A description of special aspects of the weighting scheme used in least-squares refinement. Used to describe the weighting when the value of `_refine.ls_weighting_scheme` is specified as 'calc'.

Example:

```
; Sigdel model of Konnert-Hendrickson:
Sigdel =
Afsig + Bfsig*(sin(theta)/lambda-1/6)
Afsig = 22.0, Bfsig = 150.0
at the beginning of refinement.
Afsig = 16.0, Bfsig = 60.0
at the end of refinement.
```

[refine]

**`_refine.ls_weighting_scheme`** (ucode)  
`_refine_ls_weighting_scheme` (cif\_core.dic 2.0.1)

The weighting scheme applied in the least-squares process. The standard code may be followed by a description of the weight (but see `_refine.ls_weighting_details` for a preferred approach).

The data value must be one of the following:

sigma	based on measured e.s.d.'s
unit	unit or no weights applied
calc	calculated weights applied

[refine]

**`_refine.ls_wR_factor_all`** (float)  
`_refine_ls_wR_factor_all` (cif\_core.dic 2.0.1)

Weighted residual factor  $wR$  for all reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low`.

$$wR = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2|}{\sum |wY_{\text{obs}}^2|} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed amplitude specified by `_refine.ls_structure_factor_coef`,  $Y_{\text{calc}}$  = the calculated amplitude specified by `_refine.ls_structure_factor_coef` and  $w$  = the least-squares weight; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_R_factor_all` (alternate). [refine]

**`_refine.ls_wR_factor_obs`** (float)  
`_refine_ls_wR_factor_obs` (cif\_core.dic 2.0.1)

Weighted residual factor  $wR$  for reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflns.observed_criterion`.

$$wR = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2|}{\sum |wY_{\text{obs}}^2|} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed amplitude specified by `_refine.ls_structure_factor_coef`,  $Y_{\text{calc}}$  = the calculated amplitude specified by `_refine.ls_structure_factor_coef` and  $w$  = the least-squares weight; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_R_factor_obs` (alternate). [refine]

**`_refine.ls_wR_factor_R_free`** (float)

Weighted residual factor  $wR$  for reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the test reflections (*i.e.* were excluded from the refinement) when the refinement included the calculation of a 'free'  $R$  factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`.

$$wR = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2|}{\sum |wY_{\text{obs}}^2|} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed amplitude specified by `_refine.ls_structure_factor_coef`,  $Y_{\text{calc}}$  = the calculated amplitude specified by `_refine.ls_structure_factor_coef` and  $w$  = the least-squares weight; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_R_factor_R_free` (alternate). [refine]

**`_refine.ls_wR_factor_R_work`** (float)

Weighted residual factor  $wR$  for reflections that satisfy the resolution limits established by `_refine.ls_d_res_high` and `_refine.ls_d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the working reflections (*i.e.* were included in the refinement) when the refinement included the calculation of a 'free'  $R$  factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`.

$$wR = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2|}{\sum |wY_{\text{obs}}|^2} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed amplitude specified by `_refine.ls_structure_factor_coef`,  $Y_{\text{calc}}$  = the calculated amplitude specified by `_refine.ls_structure_factor_coef` and  $w$  = the least-squares weight; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine.ls_R_factor_R_work` (alternate) [refine]

**`_refine.occupancy_max`** (float)

The maximum value for occupancy found in the coordinate set.

The permitted range is [0.0, ∞). [refine]

**`_refine.occupancy_min`** (float)

The minimum value for occupancy found in the coordinate set.

The permitted range is [0.0, ∞). [refine]

**`_refine.overall_FOM_free_R_set`** (float)

`_refine.ebi_overall_FOM_free_Rset` (ebi\_extensions 1.0)

Average figure of merit of phases of reflections not included in the refinement. This value is derived from the likelihood function

$$\text{FOM} = I_1(X)/I_0(X),$$

where  $I_0, I_1$  = zero- and first-order modified Bessel functions of the first kind,  $X = \sigma_A |E_o| |E_c| / \text{SIGMA}$ ,  $E_o, E_c$  = normalized observed and calculated structure factors,  $\sigma_A = \langle \cos 2\pi s \delta_x \rangle \sqrt{(\Sigma_P / \Sigma_N)}$  estimated using maximum likelihood,  $\Sigma_P = \sum_{\text{atoms in model}} f^2$ ,  $\Sigma_N = \sum_{\text{atoms in crystal}} f^2$ ,  $f$  = form factor of atoms,  $\delta_x$  = expected error,  $\text{SIGMA} = (\sigma_{\{E;\text{exp}\}})^2 + \varepsilon [1 - (\sigma_A)^2]$ ,  $\sigma_{\{E;\text{exp}\}}$  = uncertainties of normalized observed structure factors and  $\varepsilon$  = multiplicity of the diffracting plane.

Reference: Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst. D53*, 240–255.

[refine]

**`_refine.overall_FOM_work_R_set`** (float)

`_refine.ebi_overall_FOM_work_Rset` (ebi\_extensions 1.0)

Average figure of merit of phases of reflections included in the refinement. This value is derived from the likelihood function

$$\text{FOM} = I_1(X)/I_0(X),$$

where  $I_0, I_1$  = zero- and first-order modified Bessel functions of the first kind,  $X = \sigma_A |E_o| |E_c| / \text{SIGMA}$ ,  $E_o, E_c$  = normalized observed and calculated structure factors,  $\sigma_A = \langle \cos 2\pi s \delta_x \rangle \sqrt{(\Sigma_P / \Sigma_N)}$  estimated using maximum likelihood,  $\Sigma_P = \sum_{\text{atoms in model}} f^2$ ,  $\Sigma_N = \sum_{\text{atoms in crystal}} f^2$ ,  $f$  = form factor of atoms,  $\delta_x$  = expected error,  $\text{SIGMA} = (\sigma_{\{E;\text{exp}\}})^2 + \varepsilon [1 - (\sigma_A)^2]$ ,  $\sigma_{\{E;\text{exp}\}}$  = uncertainties of normalized observed structure factors and  $\varepsilon$  = multiplicity of the diffracting plane.

Reference: Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst. D53*, 240–255.

[refine]

**`_refine.overall_SU_B`** (float)

`_refine.ebi_Overall_ESU_B` (ebi\_extensions 1.0)

The overall standard uncertainty (estimated standard deviation) of the displacement parameters based on a maximum-likelihood residual. The overall standard uncertainty  $(\sigma_B)^2$  gives an idea of the uncertainty in the  $B$  values of averagely defined atoms (atoms with  $B$  values equal to the average  $B$  value).

$$(\sigma_B)^2 = \frac{8N_a}{\sum_i [1/\Sigma - (E_o)^2(1 - m^2)](\text{SUM\_AS})s^4},$$

where  $\text{SUM\_AS} = (\sigma_A)^2 / \Sigma^2$ ,  $N_a$  = number of atoms,  $\Sigma = (\sigma_{\{E;\text{exp}\}})^2 + \varepsilon [1 - (\sigma_A)^2]$ ,  $E_o$  = normalized structure factors,  $\sigma_{\{E;\text{exp}\}}$  = experimental uncertainties of normalized structure factors,  $\sigma_A = \langle \cos 2\pi s \delta_x \rangle \sqrt{(\Sigma_P / \Sigma_N)}$  estimated using maximum likelihood,  $\Sigma_P = \sum_{\text{atoms in model}} f^2$ ,  $\Sigma_N = \sum_{\text{atoms in crystal}} f^2$ ,  $f$  = form factor of atoms,  $\delta_x$  = expected error,  $m$  = figure of merit of phases of reflections included in the summation,  $s$  = reciprocal-space vector and  $\varepsilon$  = multiplicity of the diffracting plane; the summation is over all reflections included in refinement.

References:  $\sigma_A$  estimation: Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst. D53*, 240–255. SU ML estimation: Murshudov, G. N. & Dodson, E. J. (1997). *CCP4 Newsletter on Protein Crystallography*, No. 33, January 1997, pp. 31–39. (<http://www.ccp4.ac.uk/newsletters/newsletter33/murshudov.html>)

[refine]

**`_refine.overall_SU_ML`** (float)

`_refine.ebi_Overall_ESU_ML` (ebi\_extensions 1.0)

The overall standard uncertainty (estimated standard deviation) of the positional parameters based on a maximum-likelihood residual. The overall standard uncertainty  $(\sigma_X)^2$  gives an idea of the uncertainty in the position of averagely defined atoms (atoms with  $B$  values equal to the average  $B$  value).

$$(\sigma_X)^2 = \frac{3N_a}{8\pi^2 \sum_i [1/\Sigma - (E_o)^2(1 - m^2)](\text{SUM\_AS})s^2},$$

where  $\text{SUM\_AS} = (\sigma_A)^2 / \Sigma^2$ ,  $N_a$  = number of atoms,  $\Sigma = (\sigma_{\{E;\text{exp}\}})^2 + \varepsilon [1 - (\sigma_A)^2]$ ,  $E_o$  = normalized structure factors,  $\sigma_{\{E;\text{exp}\}}$  = experimental uncertainties of normalized structure factors,  $\sigma_A = \langle \cos 2\pi s \delta_x \rangle \sqrt{(\Sigma_P / \Sigma_N)}$  estimated using maximum likelihood,  $\Sigma_P = \sum_{\text{atoms in model}} f^2$ ,  $\Sigma_N = \sum_{\text{atoms in crystal}} f^2$ ,  $f$  = form factor of atoms,  $\delta_x$  = expected error,  $m$  = figure of merit of phases of reflections included in the summation,  $s$  = reciprocal-space vector and  $\varepsilon$  = multiplicity of the diffracting plane; the summation is over all reflections included in refinement.

References:  $\sigma_A$  estimation: Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst. D53*, 240–255. SU ML estimation: Murshudov, G. N. & Dodson, E. J. (1997). *CCP4 Newsletter on Protein Crystallography*, No. 33, January 1997, pp. 31–39. (<http://www.ccp4.ac.uk/newsletters/newsletter33/murshudov.html>)

[refine]

**`_refine.overall_SU_R_Cruickshank_DPI`** (float)  
**`_refine.ebi_Overall_ESU_R_Cruickshanks_DPI`** (*ebi\_extensions 1.0*)

The overall standard uncertainty (estimated standard deviation) of the displacement parameters based on the crystallographic  $R$  value, expressed in a formalism known as the dispersion precision indicator (DPI). The overall standard uncertainty ( $\sigma_B$ ) gives an idea of the uncertainty in the  $B$  values of averagely defined atoms (atoms with  $B$  values equal to the average  $B$  value).

$$(\sigma_B)^2 = 0.65 \frac{N_a}{(N_o - N_p)} (R_{\text{value}})^2 (D_{\text{min}})^2 C^{(-2/3)},$$

where  $N_a$  = number of atoms,  $N_o$  = number of reflections included in refinement,  $N_p$  = number of refined parameters,  $R_{\text{value}}$  = conventional crystallographic  $R$  value,  $D_{\text{min}}$  = maximum resolution and  $C$  = completeness of data.

References: Cruickshank, D. W. J. (1999). *Acta Cryst. D55*, 583–601; Murshudov, G. N. & Dodson, E. J. (1997). *CCP4 Newsletter on Protein Crystallography*, No. 33, January 1997, pp. 31–39. (<http://www.ccp4.ac.uk/newsletters/newsletter33/murshudov.html>)

[refine]

**`_refine.overall_SU_R_free`** (float)  
**`_refine.ebi_Overall_ESU_Rfree`** (*ebi\_extensions 1.0*)

The overall standard uncertainty (estimated standard deviation) of the displacement parameters based on the free  $R$  value. The overall standard uncertainty gives an idea of the uncertainty in the  $B$  values of averagely defined atoms (atoms with  $B$  values equal to the average  $B$  value).

$$(\sigma_B)^2 = 0.65 \frac{N_a}{N_o} (R_{\text{free}})^2 (D_{\text{min}})^2 C^{(-2/3)},$$

where  $N_a$  = number of atoms,  $N_o$  = number of reflections included in the refinement,  $R_{\text{free}}$  = conventional free crystallographic  $R$  value calculated using the reflections not included in the refinement,  $D_{\text{min}}$  = maximum resolution and  $C$  = completeness of data.

References: Cruickshank, D. W. J. (1999). *Acta Cryst. D55*, 583–601; Murshudov, G. N. & Dodson, E. J. (1997). *CCP4 Newsletter on Protein Crystallography*, No. 33, January 1997, pp. 31–39. (<http://www.ccp4.ac.uk/newsletters/newsletter33/murshudov.html>)

[refine]

**`_refine.solvent_model_details`** (text)  
 Special aspects of the solvent model used during refinement.

[refine]

**`_refine.solvent_model_param_bsol`** (float)

The value of the  $B_{\text{SOL}}$  solvent-model parameter describing the average isotropic displacement parameter of disordered solvent atoms. This is one of the two parameters (the other is `_refine.solvent_model_param_ksol`) in Tronrud's method of modelling the contribution of bulk solvent to the scattering. The standard scale factor is modified according to the expression

$$k_0 \exp(-B_0 s^2) [1 - K_{\text{SOL}} \exp(-B_{\text{SOL}} s^2)],$$

where  $k_0$  and  $B_0$  are the scale factors for the protein.

Reference: Tronrud, D. E. (1997). *Methods Enzymol.* **277**, 243–268.

[refine]

**`_refine.solvent_model_param_ksol`** (float)

The value of the  $K_{\text{SOL}}$  solvent-model parameter describing the ratio of the electron density in the bulk solvent to the electron density in the molecular solute. This is one of the two parameters (the other is `_refine.solvent_model_param_bsol`) in Tronrud's method of modelling the contribution of bulk solvent to the scattering. The standard scale factor is modified according to the expression

$$k_0 \exp(-B_0 s^2) [1 - K_{\text{SOL}} \exp(-B_{\text{SOL}} s^2)],$$

where  $k_0$  and  $B_0$  are the scale factors for the protein.

Reference: Tronrud, D. E. (1997). *Methods Enzymol.* **277**, 243–268.

[refine]

### REFINE\_ANALYZE

Data items in the REFINE\_ANALYZE category record details about the refined structure that are often used to analyze the refinement and assess its quality. A given computer program may or may not produce values corresponding to these data names.

Category group(s): **`inclusive_group`**

**`refine_group`**

Category key(s): **`_refine_analyze.entry_id`**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _refine_analyze.entry_id
  _refine_analyze.Luzzati_coordinate_error_obs
  _refine_analyze.Luzzati_d_res_low_obs
    5HVP 0.056 2.51
```

\* **`_refine_analyze.entry_id`**

This data item is a pointer to `_entry.id` in the ENTRY category.

[refine\_analyze]

**`_refine_analyze.Luzzati_coordinate_error_free`** (float)

The estimated coordinate error obtained from the plot of the  $R$  value versus  $\sin(\theta)/\lambda$  for the reflections treated as a test set during refinement.

Reference: Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

[refine\_analyze]

**`_refine_analyze.Luzzati_coordinate_error_obs`** (float)

The estimated coordinate error obtained from the plot of the  $R$  value versus  $\sin(\theta)/\lambda$  for reflections classified as observed.

Reference: Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

[refine\_analyze]

**`_refine_analyze.Luzzati_d_res_low_free`** (float)

The value of the low-resolution cutoff used in constructing the Luzzati plot for reflections treated as a test set during refinement.

Reference: Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

[refine\_analyze]

**`_refine_analyze.Luzzati_d_res_low_obs`** (float)

The value of the low-resolution cutoff used in constructing the Luzzati plot for reflections classified as observed.

Reference: Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

[refine\_analyze]

**`_refine_analyze.Luzzati_sigma_a_free`** (float)  
 The value of  $\sigma_a$  used in constructing the Luzzati plot for the reflections treated as a test set during refinement. Details of the estimation of  $\sigma_a$  can be specified in `_refine_analyze.Luzzati_sigma_a_free_details`.

Reference: Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

[refine\_analyze]

**`_refine_analyze.Luzzati_sigma_a_free_details`** (text)

Details of the estimation of  $\sigma_a$  for the reflections treated as a test set during refinement.

Reference: Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

[refine\_analyze]

**`_refine_analyze.Luzzati_sigma_a_obs`** (float)

The value of  $\sigma_a$  used in constructing the Luzzati plot for reflections classified as observed. Details of the estimation of  $\sigma_a$  can be specified in `_refine_analyze.Luzzati_sigma_a_obs_details`.

Reference: Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

[refine\_analyze]

**`_refine_analyze.Luzzati_sigma_a_obs_details`** (text)

Special aspects of the estimation of  $\sigma_a$  for the reflections classified as observed.

Reference: Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.

[refine\_analyze]

**`_refine_analyze.number_disordered_residues`** (float)

The number of discretely disordered residues in the refined model.

[refine\_analyze]

**`_refine_analyze.occupancy_sum_hydrogen`** (float)

The sum of the occupancies of the hydrogen atoms in the refined model.

[refine\_analyze]

**`_refine_analyze.occupancy_sum_non_hydrogen`** (float)

The sum of the occupancies of the non-hydrogen atoms in the refined model.

[refine\_analyze]

**`_refine_analyze.RG_d_res_high`** (float)

`_refine_analyze.ebi_RG_d_res_high` (ebi\_extensions 1.0)

The value of the high-resolution cutoff in ångströms used in the calculation of the Hamilton generalized  $R$  factor ( $R_G$ ) stored in `_refine_analyze.RG_work` and `_refine_analyze.RG_free`.

Reference: Hamilton, W. C. (1965). *Acta Cryst.* **18**, 502–510.

The permitted range is [0.0, ∞).

[refine\_analyze]

**`_refine_analyze.RG_d_res_low`** (float)

`_refine_analyze.ebi_RG_d_res_low` (ebi\_extensions 1.0)

The value of the low-resolution cutoff in ångströms used in the calculation of the Hamilton generalized  $R$  factor ( $R_G$ ) stored in `_refine_analyze.RG_work` and `_refine_analyze.RG_free`.

Reference: Hamilton, W. C. (1965). *Acta Cryst.* **18**, 502–510.

The permitted range is [0.0, ∞).

[refine\_analyze]

**`_refine_analyze.RG_free`** (float)

`_refine_analyze.ebi_RG_free` (ebi\_extensions 1.0)

The Hamilton generalized  $R$  factor for all reflections that satisfy the resolution limits established by `_refine_analyze.RG_d_res_high` and `_refine_analyze.RG_d_res_low` for the free  $R$  set of reflections that were excluded from the refinement.

$$R_G = \sqrt{\frac{\sum_i \sum_j w_{i,j} (|F_{\text{obs}}|_i - G|F_{\text{calc}}|_i) (|F_{\text{obs}}|_j - G|F_{\text{calc}}|_j)}{\sum_i \sum_j w_{i,j} |F_{\text{obs}}|_i |F_{\text{obs}}|_j}}$$

where  $|F_{\text{obs}}|$  = the observed structure-factor amplitudes,  $|F_{\text{calc}}|$  = the calculated structure-factor amplitudes,  $G$  = the scale factor which puts  $|F_{\text{calc}}|$  on the same scale as  $|F_{\text{obs}}|$  and  $w_{i,j}$  = the weight for the combination of the reflections  $i$  and  $j$ ;  $\sum_i$  and  $\sum_j$  are taken over the specified reflections.

When the covariance of the amplitudes of reflection  $i$  and reflection  $j$  is zero (*i.e.* the reflections are independent),  $w_{i,i}$  can be redefined as  $w_i$  and the nested sums collapsed into one sum:

$$R_G = \sqrt{\frac{\sum_i w_i (|F_{\text{obs}}|_i - G|F_{\text{calc}}|_i)^2}{\sum_i w_i |F_{\text{obs}}|_i^2}}$$

Reference: Hamilton, W. C. (1965). *Acta Cryst.* **18**, 502–510.

The permitted range is [0.0, ∞).

[refine\_analyze]

**`_refine_analyze.RG_free_work_ratio`** (float)

`_refine_analyze.ebi_RG_work_free_ratio` (ebi\_extensions 1.0)

The observed ratio of  $R_{G\text{free}}$  to  $R_{G\text{work}}$ . The expected  $R_G$  ratio is the value that should be achievable at the end of a structure refinement when only random uncorrelated errors exist in the data and the model provided that the observations are properly weighted. When compared with the observed  $R_G$  ratio, it may indicate that a structure has not reached convergence or a model has been over-refined with no corresponding improvement in the model.

In an unrestrained refinement, the ratio of  $R_{G\text{free}}$  to  $R_{G\text{work}}$  with only random uncorrelated errors at convergence depends only on the number of reflections and the number of parameters according to

$$\sqrt{(f + m)/(f - m)},$$

where  $f$  = the number of included structure amplitudes and target distances and  $m$  = the number of parameters being refined.

In the restrained case,  $R_{G\text{free}}$  is calculated from a random selection of residuals including both the structure amplitudes and the restraints. When restraints are included in the refinement, the  $R_G$  ratio requires a term for the contribution to the minimized residual at convergence,  $D_{\text{restr}}$ , due to those restraints:

$$D_{\text{restr}} = r - \sum [w_i \cdot (a_i)^t \cdot (H)^{-1} a_i],$$

where  $r$  is the number of geometrical, displacement-parameter and other restraints,  $H$  is the  $(m, m)$  normal matrix given by  $A^t \cdot W \cdot A$ ,  $W$  is the  $(n, n)$  symmetric weight matrix of the included observations,  $A$  is the least-squares design matrix of derivatives of order  $(n, m)$  and  $a_i$  is the  $i$ th row of  $A$ . Then the expected  $R_G$  ratio becomes

$$\sqrt{[f + (m - r + D_{\text{restr}})]/[f - (m - r + D_{\text{restr}})]}$$

There is no data name for the expected value of  $R_{G\text{free}}/R_{G\text{work}}$  yet.

Reference: Tickle, I. J., Laskowski, R. A. & Moss, D. S. (1998). *Acta Cryst.* **D54**, 547–557.

The permitted range is [0.0, ∞).

[refine\_analyze]

**`_refine_analyze.RG_work`** (float)  
**`_refine_analyze.ebi_RG_work`** (ebi\_extensions 1.0)

The Hamilton generalized  $R$  factor for all reflections that satisfy the resolution limits established by `_refine_analyze.RG_d_res_high` and `_refine_analyze.RG_d_res_low` and for those reflections included in the working set when a free  $R$  set of reflections is omitted from the refinement.

$$R_G = \sqrt{\frac{\sum_i \sum_j w_{i,j} (|F_{\text{obs}}|_i - G|F_{\text{calc}}|_i) (|F_{\text{obs}}|_j - G|F_{\text{calc}}|_j)}{\sum_i \sum_j w_{i,j} |F_{\text{obs}}|_i |F_{\text{obs}}|_j}}$$

where  $|F_{\text{obs}}|$  = the observed structure-factor amplitudes,  $|F_{\text{calc}}|$  = the calculated structure-factor amplitudes,  $G$  = the scale factor which puts  $|F_{\text{calc}}|$  on the same scale as  $|F_{\text{obs}}|$  and  $w_{i,j}$  = the weight for the combination of the reflections  $i$  and  $j$ ;  $\sum_i$  and  $\sum_j$  are taken over the specified reflections.

When the covariance of the amplitudes of reflection  $i$  and reflection  $j$  is zero (i.e. the reflections are independent),  $w_{i,j}$  can be redefined as  $w_i$  and the nested sums collapsed into one sum:

$$R_G = \sqrt{\frac{\sum_i w_i (|F_{\text{obs}}|_i - G|F_{\text{calc}}|_i)^2}{\sum_i w_i |F_{\text{obs}}|_i^2}}$$

Reference: Hamilton, W. C. (1965). *Acta Cryst.* **18**, 502–510.

The permitted range is [0.0, ∞).

[`refine_analyze`]

### REFINE\_B\_ISO

Data items in the `REFINE_B_ISO` category record details about the treatment of isotropic  $B$  factors (displacement parameters) during refinement.

Category group(s): `inclusive_group`  
`refine_group`  
 Category key(s): `_refine_B_iso.class`

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
loop_
  _refine_B_iso.class
  _refine_B_iso.treatment
  'protein' isotropic
  'solvent' isotropic
  'inhibitor' isotropic
```

\* `_refine_B_iso.class` (text)

A class of atoms treated similarly for isotropic  $B$ -factor (displacement-parameter) refinement.

Examples: 'all', 'protein', 'solvent', 'sugar-phosphate backbone'.

[`refine_B_iso`]

`_refine_B_iso.details` (text)

A description of special aspects of the isotropic  $B$ -factor (displacement-parameter) refinement for the class of atoms described in `_refine_B_iso.class`.

Example:

```
; The temperature factors of atoms in the side chain of Arg 92
were held fixed due to unstable behavior in refinement.
```

[`refine_B_iso`]

`_refine_B_iso.treatment` (ucode)

The treatment of isotropic  $B$ -factor (displacement-parameter) refinement for a class of atoms defined in `_refine_B_iso.class`.

The data value must be one of the following:

```
fixed
isotropic
anisotropic
```

[`refine_B_iso`]

`_refine_B_iso.value` (float)

The value of the isotropic  $B$  factor (displacement parameter) assigned to a class of atoms defined in `_refine_B_iso.class`. Meaningful only for atoms with fixed isotropic  $B$  factors.

[`refine_B_iso`]

### REFINE\_FUNCT\_MINIMIZED

Data items in the `REFINE_FUNCT_MINIMIZED` category record details about the individual terms of the function minimized during refinement.

Category group(s): `inclusive_group`  
`refine_group`  
 Category key(s): `_refine_func_minimized.type`

Example 1 – based on RESTRAIN refinement for the CCP4 test data set toxd.

```
loop_
  _refine_func_minimized.type
  _refine_func_minimized.number_terms
  _refine_func_minimized.residual
  'sum(W*Delta(Amplitude)^2'          3009      1621.3
  'sum(W*Delta(Plane+Rigid)^2'        85         56.68
  'sum(W*Delta(Distance)^2'          1219      163.59
  'sum(W*Delta(U-tempfactors)^2'      1192      69.338
```

`_refine_func_minimized.number_terms` (int)

`_ebi_refine_func_minimized.NumTerms` (ebi\_extensions 1.0)

The number of observations in this term. For example, if the term is a residual of the X-ray intensities, this item would contain the number of reflections used in the refinement.

The permitted range is [0, ∞).

[`refine_func_minimized`]

`_refine_func_minimized.residual` (float)

`_ebi_refine_func_minimized.Residual` (ebi\_extensions 1.0)

The residual for this term of the function that was minimized during the refinement.

The permitted range is [0.0, ∞).

[`refine_func_minimized`]

\* `_refine_func_minimized.type` (line)

`_ebi_refine_func_minimized.type` (ebi\_extensions 1.0)

The type of the function being minimized.

[`refine_func_minimized`]

`_refine_func_minimized.weight` (float)

`_ebi_refine_func_minimized.weight` (ebi\_extensions 1.0)

The weight applied to this term of the function that was minimized during the refinement.

[`refine_func_minimized`]

### REFINE\_HIST

Data items in the `REFINE_HIST` category record details about the steps during the refinement of the structure. These data items are not meant to be as thorough a description of the refinement as is provided for the final model in other categories; rather, these data items provide a mechanism for sketching out the progress of the refinement, supported by a small set of representative statistics.

Category group(s): `inclusive_group`  
`refine_group`  
 Category key(s): `_refine_hist.cycle_id`

*Example 1 – based on laboratory records for the collagen-like peptide [(POG)<sub>5</sub>]<sub>3</sub> EKG (POG)<sub>5</sub>]<sub>3</sub>.*

```

_refine_hist.cycle_id          C134
_refine_hist.d_res_high       1.85
_refine_hist.d_res_low       20.0
_refine_hist.number_atoms_solvent 217
_refine_hist.number_atoms_total 808
_refine_hist.number_reflns_all 6174
_refine_hist.number_reflns_obs 4886
_refine_hist.number_reflns_R_free 476
_refine_hist.number_reflns_R_work 4410
_refine_hist.R_factor_all     .265
_refine_hist.R_factor_obs     .195
_refine_hist.R_factor_R_free  .274
_refine_hist.R_factor_R_work  .160
_refine_hist.details
; Add majority of solvent molecules. B factors refined by
  group. Continued to remove misplaced water molecules.
;

```

\* **\_refine\_hist.cycle\_id** (code)  
 The value of `_refine_hist.cycle_id` must uniquely identify a record in the `REFINE_HIST` list. Note that this item need not be a number; it can be any unique identifier.

[refine\_hist]

\* **\_refine\_hist.d\_res\_high** (float)  
 The lowest value for the interplanar spacings for the reflection data for this cycle of refinement. This is called the highest resolution.  
 The permitted range is [0.0, ∞).

[refine\_hist]

\* **\_refine\_hist.d\_res\_low** (float)  
 The highest value for the interplanar spacings for the reflection data for this cycle of refinement. This is called the lowest resolution.

The permitted range is [0.0, ∞).

[refine\_hist]

**\_refine\_hist.details** (text)  
 A description of special aspects of this cycle of the refinement process.

Example:

```

; Residues 13-17 fit and added to model; substantial
rebuilding of loop containing residues 43-48; addition of
first atoms to solvent model; ten cycles of Prolsq refinement.
;

```

[refine\_hist]

**\_refine\_hist.number\_atoms\_solvent** (int)  
 The number of solvent atoms that were included in the model at this cycle of the refinement.

The permitted range is [0, ∞).

[refine\_hist]

**\_refine\_hist.number\_atoms\_total** (int)  
 The total number of atoms that were included in the model at this cycle of the refinement.

The permitted range is [0, ∞).

[refine\_hist]

**\_refine\_hist.number\_reflns\_all** (int)  
 The number of reflections that satisfy the resolution limits established by `_refine_hist.d_res_high` and `_refine_hist.d_res_low`.

The permitted range is [0, ∞).

[refine\_hist]

**\_refine\_hist.number\_reflns\_obs** (int)  
 The number of reflections that satisfy the resolution limits established by `_refine_hist.d_res_high` and `_refine_hist.d_res_low` and the observation criterion established by `_reflns.observed_criterion`.

The permitted range is [0, ∞).

[refine\_hist]

**\_refine\_hist.number\_reflns\_R\_free** (int)  
 The number of reflections that satisfy the resolution limits established by `_refine_hist.d_res_high` and `_refine_hist.d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the test reflections (*i.e.* were excluded from the refinement) when the refinement included the calculation of a ‘free’ *R* factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`.

The permitted range is [0, ∞).

[refine\_hist]

**\_refine\_hist.number\_reflns\_R\_work** (int)  
 The number of reflections that satisfy the resolution limits established by `_refine_hist.d_res_high` and `_refine_hist.d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the working reflections (*i.e.* were included in the refinement) when the refinement included the calculation of a ‘free’ *R* factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`.

The permitted range is [0, ∞).

[refine\_hist]

**\_refine\_hist.R\_factor\_all** (float)  
 Residual factor *R* for reflections that satisfy the resolution limits established by `_refine_hist.d_res_high` and `_refine_hist.d_res_low`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|},$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes and  $F_{\text{calc}}$  = the calculated structure-factor amplitudes; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

[refine\_hist]

**\_refine\_hist.R\_factor\_obs** (float)  
 Residual factor *R* for reflections that satisfy the resolution limits established by `_refine_hist.d_res_high` and `_refine_hist.d_res_low` and the observation criterion established by `_reflns.observed_criterion`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|},$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes and  $F_{\text{calc}}$  = the calculated structure-factor amplitudes; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

[refine\_hist]

**\_refine\_hist.R\_factor\_R\_free** (float)  
 Residual factor *R* for reflections that satisfy the resolution limits established by `_refine_hist.d_res_high` and `_refine_hist.d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the test reflections (*i.e.* were excluded from the refinement) when the refinement included the calculation of a ‘free’ *R* factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|},$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes and  $F_{\text{calc}}$  = the calculated structure-factor amplitudes; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

[refine\_hist]

**`_refine_hist.R_factor_R_work`** (float)

Residual factor  $R$  for reflections that satisfy the resolution limits established by `_refine_hist.d_res_high` and `_refine_hist.d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the working reflections (*i.e.* were included in the refinement) when the refinement included the calculation of a ‘free’  $R$  factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|}$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes and  $F_{\text{calc}}$  = the calculated structure-factor amplitudes; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

[refine\_hist]

### REFINE\_LS\_CLASS

Data items in the REFINE\_LS\_CLASS category record details about the reflections used for the structure refinement for each reflection class separately.

Category key(s): `_refine_ls_class.code`

*Example 1 – data for a modulated structure from van Smaalen [J. Phys. Condens. Matter (1991), 3, 1247–1263].*

```
loop_
  _refine_ls_class.R_factor_gt
  _refine_ls_class.code
    0.057 'Main'
    0.074 'Com'
    0.064 'NbRefls'
    0.046 'LaRefls'
    0.112 'Sat1'
    0.177 'Sat2'
```

\***`_refine_ls_class.code`** (code)

`_refine_ls_class.code` (cif\_core.dic 2.3)

The code identifying a certain reflection class. This code must match a `_reflns_class.code`.

Examples: ‘1’, ‘m1’, ‘s2’.

[refine\_ls\_class]

**`_refine_ls_class.d_res_high`** (float)

`_refine_ls_class.d_res_high` (cif\_core.dic 2.3)

For each reflection class, the lowest value in ångströms for the interplanar spacings for the reflections used in the refinement. This is called the highest resolution.

The permitted range is [0.0, ∞).

[refine\_ls\_class]

**`_refine_ls_class.d_res_low`** (float)

`_refine_ls_class.d_res_low` (cif\_core.dic 2.3)

For each reflection class, the highest value in ångströms for the interplanar spacings for the reflections used in the refinement. This is called the lowest resolution.

The permitted range is [0.0, ∞).

[refine\_ls\_class]

**`_refine_ls_class.R_factor_all`** (float)

`_refine_ls_class.R_factor_all` (cif\_core.dic 2.3)

For each reflection class, the residual factor for all reflections satisfying the resolution limits established by `_refine_ls_class.d`

`res_high` and `_refine_ls_class.d_res_low`. This is the conventional  $R$  factor. See also the definition of `_refine_ls_class.wR_factor_all`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|}$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the reflections of this class.

The permitted range is [0.0, ∞).

[refine\_ls\_class]

**`_refine_ls_class.R_factor_gt`** (float)

`_refine_ls_class.R_factor_gt` (cif\_core.dic 2.3)

For each reflection class, the residual factor for significantly intense reflections (see `_reflns.threshold_expression`) included in the refinement. The reflections also satisfy the resolution limits established by `_refine_ls_class.d_res_high` and `_refine_ls_class.d_res_low`. This is the conventional  $R$  factor. See also the definition of `_refine_ls_class.wR_factor_all`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|}$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the reflections of this class.

The permitted range is [0.0, ∞).

[refine\_ls\_class]

**`_refine_ls_class.R_Fsqd_factor`** (float)

`_refine_ls_class.R_Fsqd_factor` (cif\_core.dic 2.3)

For each reflection class, the residual factor  $R(F^2)$  calculated on the squared amplitudes of the observed and calculated structure factors for the reflections judged significantly intense (*i.e.* satisfying the threshold specified by `_reflns.threshold_expression`) and included in the refinement. The reflections also satisfy the resolution limits established by `_refine_ls_class.d_res_high` and `_refine_ls_class.d_res_low`.

$$R(F^2) = \frac{\sum |F_{\text{obs}}^2 - F_{\text{calc}}^2|}{\sum |F_{\text{obs}}^2|}$$

where  $F_{\text{obs}}^2$  = squares of the observed structure-factor amplitudes,  $F_{\text{calc}}^2$  = squares of the calculated structure-factor amplitudes and the sum is taken over the reflections of this class.

The permitted range is [0.0, ∞).

[refine\_ls\_class]

**`_refine_ls_class.R_I_factor`** (float)

`_refine_ls_class.R_I_factor` (cif\_core.dic 2.3)

For each reflection class, the residual factor  $R(I)$  for the reflections judged significantly intense (*i.e.* satisfying the threshold specified by `_reflns.threshold_expression`) and included in the refinement. This is most often calculated in Rietveld refinements against powder data, where it is referred to as  $R_B$  or  $R_{\text{Bragg}}$ .

$$R(I) = \frac{\sum |I_{\text{obs}} - I_{\text{calc}}|}{\sum |I_{\text{obs}}|}$$

where  $I_{\text{obs}}$  = the net observed intensities,  $I_{\text{calc}}$  = the net calculated intensities and the sum is taken over the reflections of this class.

The permitted range is [0.0, ∞).

[refine\_ls\_class]

**`_refine_ls_class.wr_factor_all`** (float)  
`_refine_ls_class.wr_factor_all` (cif-core.dic 2.3)

For each reflection class, the weighted residual factor for all reflections included in the refinement. The reflections also satisfy the resolution limits established by `_refine_ls_class.d_res_high` and `_refine_ls_class.d_res_low`. See also the `_refine_ls_class.R_factor_definitions`.

$$wR = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2|}{\sum |wY_{\text{obs}}|^2} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed amplitudes specified by `_refine_ls_structure_factor_coef`,  $Y_{\text{calc}}$  = the calculated amplitudes specified by `_refine_ls_structure_factor_coef`,  $w$  = the least-squares weight and the sum is taken over the reflections of this class.

The permitted range is [0.0, ∞).

[refine\_ls\_class]

### REFINE\_LS\_RESTR

Data items in the REFINE\_LS\_RESTR category record details about the restraints applied to various classes of parameters during the least-squares refinement.

Category group(s): `inclusive_group`  
`refine_group`

Category key(s): `_refine_ls_restr.type`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _refine_ls_restr.type
  _refine_ls_restr.dev_ideal_target
  _refine_ls_restr.dev_ideal
  _refine_ls_restr.number
  _refine_ls_restr.criterion
  _refine_ls_restr.rejects
  'p_bond_d'      0.020  0.018  1654  '> 2\s'  22
  'p_angle_d'    0.030  0.038  2246  '> 2\s'  139
  'p_planar_d'   0.040  0.043  498   '> 2\s'  21
  'p_planar'     0.020  0.015  270   '> 2\s'  1
  'p_chiral'     0.150  0.177  278   '> 2\s'  2
  'p_singtor_nbd' 0.500  0.216  582   '> 2\s'  0
  'p_multtor_nbd' 0.500  0.207  419   '> 2\s'  0
  'p_xyhbond_nbd' 0.500  0.245  149   '> 2\s'  0
  'p_planar_tor' 3.0    2.6    203   '> 2\s'  9
  'p_staggered_tor' 15.0  17.4  298   '> 2\s'  31
  'p_orthonormal_tor' 20.0  18.1  12    '> 2\s'  1
```

**`_refine_ls_restr.criterion`** (text)

A criterion used to define a parameter value that deviates significantly from its ideal value in the model obtained by restrained least-squares refinement.

Example: '> 3\s'.

[refine\_ls\_restr]

**`_refine_ls_restr.dev_ideal`** (float)

For the given parameter type, the root-mean-square deviation between the ideal values used as restraints in the least-squares refinement and the values obtained by refinement. For instance, bond distances may deviate by 0.018 Å (r.m.s.) from ideal values in the current model.

The permitted range is [0.0, ∞).

[refine\_ls\_restr]

**`_refine_ls_restr.dev_ideal_target`** (float)

For the given parameter type, the target root-mean-square deviation between the ideal values used as restraints in the least-squares refinement and the values obtained by refinement.

The permitted range is [0.0, ∞).

[refine\_ls\_restr]

**`_refine_ls_restr.number`** (int)

The number of parameters of this type subjected to restraint in least-squares refinement.

The permitted range is [0, ∞).

[refine\_ls\_restr]

**`_refine_ls_restr.rejects`** (int)

The number of parameters of this type that deviate from ideal values by more than the amount defined in `_refine_ls_restr.criterion` in the model obtained by restrained least-squares refinement.

The permitted range is [0, ∞).

[refine\_ls\_restr]

**\* `_refine_ls_restr.type`** (line)

The type of the parameter being restrained. Explicit sets of data values are provided for the programs *PROTIN/PROLSQ* (beginning with *p\_*) and *RESTRAIN* (beginning with *RESTRAIN\_*). As computer programs change, these data values are given as examples, not as an enumeration list. Computer programs that convert a data block to a refinement table will expect the exact form of the data values given here to be used.

The following item(s) have an equivalent role in their respective categories:

**`_refine_ls_restr.type.type`**.

Examples: 'p\_bond\_d' (bond distance), 'p\_angle\_d' (bond angle expressed as a distance), 'p\_planar\_d' (planar 1,4 distance), 'p\_xhbond\_d' (X—H bond distance), 'p\_xhangle\_d' (X—H bond angle expressed as a distance), 'p\_hydrog\_d' (hydrogen distance), 'p\_special\_d' (special distance), 'p\_planar' (planes), 'p\_chiral' (chiral centres), 'p\_singtor\_nbd' (single-torsion non-bonded contact), 'p\_multtor\_nbd' (multiple-torsion non-bonded contact), 'p\_xyhbond\_nbd' (possible X··Y hydrogen bond), 'p\_xyhbond\_nbd' (possible X—H··Y hydrogen bond), 'p\_special\_tor' (special torsion angle), 'p\_planar\_tor' (planar torsion angle), 'p\_staggered\_tor' (staggered torsion angle), 'p\_orthonormal\_tor' (orthonormal torsion angle), 'p\_mcbond\_it' (main-chain bond isotropic displacement parameter), 'p\_mcbond\_it' (main-chain angle isotropic displacement parameter), 'p\_scbond\_it' (side-chain bond isotropic displacement parameter), 'p\_scangle\_it' (side-chain angle isotropic displacement parameter), 'p\_xhbond\_it' (X—H bond isotropic displacement parameter), 'p\_xhangle\_it' (X—H angle isotropic displacement parameter), 'p\_special\_it' (special isotropic displacement parameter), 'RESTRAIN\_Distances < 2.12' (the root-mean-square deviation of the difference between the values calculated from the structures used to compile the restraints dictionary parameters and the dictionary values themselves in the distance range less than 2.12 Å), 'RESTRAIN\_Distances 2.12 < D < 2.625' (the root-mean-square deviation of the difference between the values calculated from the structures used to compile the restraints dictionary parameters and the dictionary values themselves in the distance range 2.12–2.625 Å), 'RESTRAIN\_Distances > 2.625' (the root-mean-square deviation of the difference between the values calculated from the structures used to compile the restraints dictionary parameters and the dictionary values themselves in the distance range greater than 2.625 Å), 'RESTRAIN\_Peptide Planes' (the root-mean-square deviation of the difference between the values calculated from the structures used to compile the restraints dictionary parameters and the dictionary values themselves for peptide planes), 'RESTRAIN\_Ring and other planes' (the root-mean-square deviation of the difference between the values calculated from the structures used to compile the restraints dictionary parameters and the dictionary values themselves for rings and planes other than peptide planes), 'RESTRAIN\_rms diffs for Uiso atoms at dist 1.2-1.4', 'RESTRAIN\_rms diffs for Uiso atoms at dist 1.4-1.6', 'RESTRAIN\_rms diffs for Uiso atoms at dist 1.8-2.0', 'RESTRAIN\_rms diffs for Uiso atoms at dist 2.0-2.2', 'RESTRAIN\_rms diffs for Uiso atoms at dist 2.2-2.4', 'RESTRAIN\_rms diffs for Uiso atoms at dist >2.4'.

[refine\_ls\_restr]

**`_refine_ls_restr.weight`** (float)

The weighting value applied to this type of restraint in the least-squares refinement.

[refine\_ls\_restr]

## REFINE\_LS\_RESTR\_NCS

Data items in the REFINE\_LS\_RESTR\_NCS category record details about the restraints applied to atom positions in domains related by noncrystallographic symmetry during least-squares refinement, and also about the deviation of the restrained atomic parameters at the end of the refinement. It is expected that these values will only be reported once for each set of restrained domains.

Category group(s): `inclusive_group`  
`refine_group`

Category key(s): `_refine_ls_restr_ncs.dom_id`

*Example 1 – based on laboratory records for the collagen-like peptide, HYP-.*

```
_refine_ls_restr_ncs.dom_id          d2
_refine_ls_restr_ncs.weight_position 300.0
_refine_ls_restr_ncs.weight_B_iso    2.0
_refine_ls_restr_ncs.rms_dev_position 0.09
_refine_ls_restr_ncs.rms_dev_B_iso   0.16
_refine_ls_restr_ncs.ncs_model_details
;
NCS restraint for pseudo-twofold symmetry between domains
d1 and d2. Position weight coefficient given in
Kcal/(mol %A^2) and isotropic B weight coefficient given
in %A^2.
;
```

\* `_refine_ls_restr_ncs.dom_id` (code)

This data item is a pointer to `_struct_ncs_dom.id` in the `STRUCT_NCS_DOM` category.

[`refine_ls_restr_ncs`]

`_refine_ls_restr_ncs.ncs_model_details` (text)

Special aspects of the manner in which noncrystallographic restraints were applied to atomic parameters in the domain specified by `_refine_ls_restr_ncs.dom_id` and equivalent atomic parameters in the domains against which it was restrained.

[`refine_ls_restr_ncs`]

`_refine_ls_restr_ncs.rms_dev_B_iso` (float)

The root-mean-square deviation in equivalent isotropic displacement parameters in the domain specified by `_refine_ls_restr_ncs.dom_id` and in the domains against which it was restrained.

The permitted range is [0.0, ∞).

[`refine_ls_restr_ncs`]

`_refine_ls_restr_ncs.rms_dev_position` (float)

The root-mean-square deviation in equivalent atom positions in the domain specified by `_refine_ls_restr_ncs.dom_id` and in the domains against which it was restrained.

The permitted range is [0.0, ∞).

[`refine_ls_restr_ncs`]

`_refine_ls_restr_ncs.weight_B_iso` (float)

The value of the weighting coefficient used in noncrystallographic symmetry restraint of isotropic displacement parameters in the domain specified by `_refine_ls_restr_ncs.dom_id` to equivalent isotropic displacement parameters in the domains against which it was restrained.

[`refine_ls_restr_ncs`]

`_refine_ls_restr_ncs.weight_position` (float)

The value of the weighting coefficient used in noncrystallographic symmetry restraint of atom positions in the domain specified by `_refine_ls_restr_ncs.dom_id` to equivalent atom positions in the domains against which it was restrained.

[`refine_ls_restr_ncs`]

## REFINE\_LS\_RESTR\_TYPE

Data items in the REFINE\_LS\_RESTR\_TYPE category record details about the restraint types used in the least-squares refinement.

Category group(s): `inclusive_group`  
`refine_group`

Category key(s): `_refine_ls_restr_type.type`

*Example 1 – based on RESTRAIN refinement for the CCP4 test data set toxd.*

```
loop_
_refine_ls_restr_type.type
_refine_ls_restr.number
_refine_ls_restr.dev_ideal
_refine_ls_restr.dev_ideal_target
'RESTRAIN Distances < 2.12'          509    0.005 0.022
'RESTRAIN Distances 2.12 < D < 2.625' 671    0.016 0.037
'RESTRAIN Distances > 2.625'          39     0.034 0.043
'RESTRAIN Peptide Planes'             59     0.002 0.010
'RESTRAIN Ring and other planes'       26     0.014 0.010
'RESTRAIN_rms diffs for Uiso atoms at dist 1.2-1.4'
  212    0.106 .
'RESTRAIN_rms diffs for Uiso atoms at dist 1.4-1.6'
  288    0.101 .
'RESTRAIN_rms diffs for Uiso atoms at dist 1.8-2.0'
   6     0.077 .
'RESTRAIN_rms diffs for Uiso atoms at dist 2.0-2.2'
  10     0.114 .
'RESTRAIN_rms diffs for Uiso atoms at dist 2.2-2.4'
  215    0.119 .
'RESTRAIN_rms diffs for Uiso atoms at dist >2.4'
  461    0.106 .
```

```
loop_
_refine_ls_restr_type.type
_refine_ls_restr_type.distance_cutoff_low
_refine_ls_restr_type.distance_cutoff_high
'RESTRAIN Distances < 2.12'          .      2.12
'RESTRAIN Distances 2.12 < D < 2.625' 2.12  2.625
'RESTRAIN Distances > 2.625'          2.625 .
'RESTRAIN Peptide Planes'             .      .
'RESTRAIN Ring and other planes'       .      .
'RESTRAIN_rms diffs for Uiso atoms at dist 1.2-1.4'
  1.2   1.4
'RESTRAIN_rms diffs for Uiso atoms at dist 1.4-1.6'
  1.4   1.6
'RESTRAIN_rms diffs for Uiso atoms at dist 1.8-2.0'
  1.8   2.0
'RESTRAIN_rms diffs for Uiso atoms at dist 2.0-2.2'
  2.0   2.2
'RESTRAIN_rms diffs for Uiso atoms at dist 2.2-2.4'
  2.2   2.4
'RESTRAIN_rms diffs for Uiso atoms at dist >2.4'
  2.4   .
```

`_refine_ls_restr_type.distance_cutoff_high` (float)

The upper limit in ångströms of the distance range applied to the current restraint type.

The permitted range is [0.0, ∞).

[`refine_ls_restr_type`]

`_refine_ls_restr_type.distance_cutoff_low` (float)

The lower limit in ångströms of the distance range applied to the current restraint type.

The permitted range is [0.0, ∞).

[`refine_ls_restr_type`]

\* `_refine_ls_restr_type.type` (line)

This data item is a pointer to `_refine_ls_restr.type` in the `REFINE_LS_RESTR` category.

[`refine_ls_restr_type`]

REFINE_LS_SHELL				
Data items in the REFINE_LS_SHELL category record details about the results of the least-squares refinement broken down into shells of resolution.				
Category group(s): <b>inclusive_group</b>				
<b>refine_group</b>				
Category key(s): <b>_refine_ls_shell.d_res_low</b>				
<b>_refine_ls_shell.d_res_high</b>				
<i>Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.</i>				
loop_				
_refine_ls_shell.d_res_low				
_refine_ls_shell.d_res_high				
_refine_ls_shell.number_reflms_obs				
_refine_ls_shell.R_factor_obs				
8.00	4.51	1226	0.196	
4.51	3.48	1679	0.146	
3.48	2.94	2014	0.160	
2.94	2.59	2147	0.182	
2.59	2.34	2127	0.193	
2.34	2.15	2061	0.203	
2.15	2.00	1647	0.188	

\* **\_refine\_ls\_shell.d\_res\_high** (float)  
 The lowest value for the interplanar spacings for the reflection data in this shell. This is called the highest resolution.  
 The permitted range is [0.0, ∞). [refine\_ls\_shell]

\* **\_refine\_ls\_shell.d\_res\_low** (float)  
 The highest value for the interplanar spacings for the reflection data in this shell. This is called the lowest resolution.  
 The permitted range is [0.0, ∞). [refine\_ls\_shell]

**\_refine\_ls\_shell.number\_reflms\_all** (int)  
 The number of reflections that satisfy the resolution limits established by **\_refine\_ls\_shell.d\_res\_high** and **\_refine\_ls\_shell.d\_res\_low**.  
 The permitted range is [0, ∞). [refine\_ls\_shell]

**\_refine\_ls\_shell.number\_reflms\_obs** (int)  
 The number of reflections that satisfy the resolution limits established by **\_refine\_ls\_shell.d\_res\_high** and **\_refine\_ls\_shell.d\_res\_low** and the observation criterion established by **\_reflms.observed\_criterion**.  
 The permitted range is [0, ∞). [refine\_ls\_shell]

**\_refine\_ls\_shell.number\_reflms\_R\_free** (int)  
 The number of reflections that satisfy the resolution limits established by **\_refine\_ls\_shell.d\_res\_high** and **\_refine\_ls\_shell.d\_res\_low** and the observation limit established by **\_reflms.observed\_criterion**, and that were used as the test reflections (*i.e.* were excluded from the refinement) when the refinement included the calculation of a ‘free’ *R* factor. Details of how reflections were assigned to the working and test sets are given in **\_reflms.R\_free\_details**.  
 The permitted range is [0, ∞). [refine\_ls\_shell]

**\_refine\_ls\_shell.number\_reflms\_R\_work** (int)  
 The number of reflections that satisfy the resolution limits established by **\_refine\_ls\_shell.d\_res\_high** and **\_refine\_ls\_shell.d\_res\_low** and the observation limit established by **\_reflms.observed\_criterion**, and that were used as the working reflections (*i.e.* were included in the refinement) when the refinement included the calculation of a ‘free’ *R* factor. Details of how reflections were assigned to the working and test sets are given in **\_reflms.R\_free\_details**.  
 The permitted range is [0, ∞). [refine\_ls\_shell]

**\_refine\_ls\_shell.percent\_reflms\_obs** (float)  
 The number of reflections that satisfy the resolution limits established by **\_refine\_ls\_shell.d\_res\_high** and **\_refine\_ls\_shell.d\_res\_low** and the observation criterion established by **\_reflms.observed\_criterion**, expressed as a percentage of the number of geometrically observable reflections that satisfy the resolution limits.  
 [refine\_ls\_shell]

**\_refine\_ls\_shell.percent\_reflms\_R\_free** (float)  
 The number of reflections that satisfy the resolution limits established by **\_refine\_ls\_shell.d\_res\_high** and **\_refine\_ls\_shell.d\_res\_low** and the observation limit established by **\_reflms.observed\_criterion**, and that were used as the test reflections (*i.e.* were excluded from the refinement) when the refinement included the calculation of a ‘free’ *R* factor, expressed as a percentage of the number of geometrically observable reflections that satisfy the reflection limits.  
 [refine\_ls\_shell]

**\_refine\_ls\_shell.R\_factor\_all** (float)  
 Residual factor *R* for reflections that satisfy the resolution limits established by **\_refine\_ls\_shell.d\_res\_high** and **\_refine\_ls\_shell.d\_res\_low**.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|}$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the specified reflections.  
 The permitted range is [0.0, ∞).  
 Related item: **\_refine\_ls\_shell.wR\_factor\_all** (alternate). [refine\_ls\_shell]

**\_refine\_ls\_shell.R\_factor\_obs** (float)  
 Residual factor *R* for reflections that satisfy the resolution limits established by **\_refine\_ls\_shell.d\_res\_high** and **\_refine\_ls\_shell.d\_res\_low** and the observation criterion established by **\_reflms.observed\_criterion**.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|}$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the specified reflections.  
 The permitted range is [0.0, ∞).  
 Related item: **\_refine\_ls\_shell.wR\_factor\_obs** (alternate). [refine\_ls\_shell]

**`_refine_ls_shell.R_factor_R_free`** (float)

Residual factor  $R$  for reflections that satisfy the resolution limits established by `_refine_ls_shell.d_res_high` and `_refine_ls_shell.d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the test reflections (*i.e.* were excluded from the refinement) when the refinement included the calculation of a ‘free’  $R$  factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|},$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related items: `_refine_ls_shell.wR_factor_R_free` (alternate),

`_refine_ls_shell.R_factor_R_free_error` (associated error).

[refine\_ls\_shell]

**`_refine_ls_shell.R_factor_R_free_error`** (float)

The estimated error in `_refine_ls_shell.R_factor_R_free`. The method used to estimate the error is described in the item `_refine_ls_R_factor_R_free_error_details`.

Related item: `_refine_ls_shell.R_factor_R_free` (associated value).

[refine\_ls\_shell]

**`_refine_ls_shell.R_factor_R_work`** (float)

Residual factor  $R$  for reflections that satisfy the resolution limits established by `_refine_ls_shell.d_res_high` and `_refine_ls_shell.d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the working reflections (*i.e.* were included in the refinement) when the refinement included the calculation of a ‘free’  $R$  factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|},$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine_ls_shell.wR_factor_R_work` (alternate).

[refine\_ls\_shell]

**`_refine_ls_shell.redundancy_reflns_all`** (float)

The ratio of the total number of observations of the reflections that satisfy the resolution limits established by `_refine_ls_shell.d_res_high` and `_refine_ls_shell.d_res_low` to the number of crystallographically unique reflections that satisfy the same limits.

[refine\_ls\_shell]

**`_refine_ls_shell.redundancy_reflns_obs`** (float)

The ratio of the total number of observations of the reflections that satisfy the resolution limits established by `_refine_ls_shell.d_res_high` and `_refine_ls_shell.d_res_low` and the observation criterion established by `_reflns.observed_criterion` to the number of crystallographically unique reflections that satisfy the same limits.

[refine\_ls\_shell]

**`_refine_ls_shell.wR_factor_all`** (float)

Weighted residual factor  $wR$  for reflections that satisfy the resolution limits established by `_refine_ls_shell.d_res_high` and `_refine_ls_shell.d_res_low`.

$$wR = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2|}{\sum |wY_{\text{obs}}^2|} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed amplitude specified by `_refine_ls_structure_factor_coef`,  $Y_{\text{calc}}$  = the calculated amplitude specified by `_refine_ls_structure_factor_coef` and  $w$  = the least-squares weight; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine_ls_shell.R_factor_all` (alternate).

[refine\_ls\_shell]

**`_refine_ls_shell.wR_factor_obs`** (float)

Weighted residual factor  $wR$  for reflections that satisfy the resolution limits established by `_refine_ls_shell.d_res_high` and `_refine_ls_shell.d_res_low` and the observation criterion established by `_reflns.observed_criterion`.

$$wR = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2|}{\sum |wY_{\text{obs}}^2|} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed amplitude specified by `_refine_ls_structure_factor_coef`,  $Y_{\text{calc}}$  = the calculated amplitude specified by `_refine_ls_structure_factor_coef` and  $w$  = the least-squares weight; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine_ls_shell.R_factor_obs` (alternate).

[refine\_ls\_shell]

**`_refine_ls_shell.wR_factor_R_free`** (float)

Weighted residual factor  $wR$  for reflections that satisfy the resolution limits established by `_refine_ls_shell.d_res_high` and `_refine_ls_shell.d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the test reflections (*i.e.* were excluded from the refinement) when the refinement included the calculation of a ‘free’  $R$  factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`.

$$wR = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2|}{\sum |wY_{\text{obs}}^2|} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed amplitude specified by `_refine_ls_structure_factor_coef`,  $Y_{\text{calc}}$  = the calculated amplitude specified by `_refine_ls_structure_factor_coef` and  $w$  = the least-squares weight; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine_ls_shell.R_factor_R_free` (alternate).

[refine\_ls\_shell]

**`_refine_ls_shell.wR_factor_R_work`** (float)  
 Weighted residual factor  $wR$  for reflections that satisfy the resolution limits established by `_refine_ls_shell.d_res_high` and `_refine_ls_shell.d_res_low` and the observation limit established by `_reflns.observed_criterion`, and that were used as the working reflections (*i.e.* were included in the refinement) when the refinement included the calculation of a 'free'  $R$  factor. Details of how reflections were assigned to the working and test sets are given in `_reflns.R_free_details`.

$$wR = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2|}{\sum |wY_{\text{obs}}|^2} \right)^{1/2},$$

where  $Y_{\text{obs}}$  = the observed amplitude specified by `_refine.ls_structure_factor_coef`,  $Y_{\text{calc}}$  = the calculated amplitude specified by `_refine.ls_structure_factor_coef` and  $w$  = the least-squares weight; the sum is taken over the specified reflections.

The permitted range is [0.0, ∞).

Related item: `_refine_ls_shell.R_factor_R_work` (alternate).

[`refine_ls_shell`]

### REFINE\_OCCUPANCY

Data items in the REFINE\_OCCUPANCY category record details about the treatment of atom occupancies during refinement.

Category group(s): `inclusive_group`

`refine_group`

Category key(s): `_refine_occupancy.class`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _refine_occupancy.class
  _refine_occupancy.treatment
  _refine_occupancy.value
  _refine_occupancy.details
  'protein'          fix 1.00 .
  'solvent'          fix 1.00 .
  'inhibitor orientation 1' fix 0.65 .
  'inhibitor orientation 2' fix 0.35
; The inhibitor binds to the enzyme in two alternative
conformations. The occupancy of each conformation was
adjusted so as to result in approximately equal mean
thermal factors for the atoms in each conformation.
;
```

\* **`_refine_occupancy.class`** (text)

The class of atoms treated similarly for occupancy refinement.

Examples: 'all', 'protein', 'solvent', 'sugar-phosphate backbone'.

[`refine_occupancy`]

**`_refine_occupancy.details`** (text)

A description of special aspects of the occupancy refinement for a class of atoms described in `_refine_occupancy.class`.

Example:

```
; The inhibitor binds to the enzyme in two alternative
conformations. The occupancy of each conformation was
adjusted so as to result in approximately equal mean thermal
factors for the atoms in each conformation.
```

[`refine_occupancy`]

**`_refine_occupancy.treatment`** (ucode)

The treatment of occupancies for a class of atoms described in `_refine_occupancy.class`.

The data value must be one of the following:

```
fix      fixed
ref      refined
```

[`refine_occupancy`]

**`_refine_occupancy.value`** (float)

The value of occupancy assigned to a class of atoms defined in `_refine_occupancy.class`. Meaningful only for atoms with fixed occupancy.

The permitted range is [0.0, 1.0].

Examples: '1.0', '0.41'.

[`refine_occupancy`]

### REFLN

Data items in the REFLN category record details about the reflection data used to determine the ATOM\_SITE data items. The REFLN data items refer to individual reflections and must be included in looped lists. The REFLNS data items specify the parameters that apply to all reflections. The REFLNS data items are not looped.

Category group(s): `inclusive_group`

`refln_group`

Category key(s): `_refln.index_h`

`_refln.index_k`

`_refln.index_l`

*Example 1 – based on data set fetod of Todres, Yanovsky, Ermekov & Struchkov [Acta Cryst. (1993), C49, 1352–1354].*

```
loop_
  _refln.index_h
  _refln.index_k
  _refln.index_l
  _refln.F_squared_calc
  _refln.F_squared_meas
  _refln.F_squared_sigma
  _refln.status
  2  0  0      85.57      58.90      1.45 o
  3  0  0    15718.18    15631.06    30.40 o
  4  0  0   55613.11   49840.09    61.86 o
  5  0  0    246.85     241.86     10.02 o
  6  0  0     82.16      69.97      1.93 o
  7  0  0    1133.62     947.79     11.78 o
  8  0  0    2558.04    2453.33    20.44 o
  9  0  0     283.88     393.66      7.79 o
 10  0  0     283.70     171.98      4.26 o
```

**`_refln.A_calc`** (float)

`_refln_A_calc` (cif\_core.dic 2.0.1)

The calculated value of structure-factor component  $A$  in electrons.

$$A = |F| \cos(\text{phase}).$$

Related item: `_refln.A_calc_au` (conversion arbitrary).

[`refln`]

**`_refln.A_calc_au`** (float)

The calculated value of structure-factor component  $A$  in arbitrary units.

$$A = |F| \cos(\text{phase}).$$

Related item: `_refln.A_calc` (conversion arbitrary).

[`refln`]

**`_refln.A_meas`** (float)

`_refln_A_meas` (cif\_core.dic 2.0.1)

The measured value of structure-factor component  $A$  in electrons.

$$A = |F| \cos(\text{phase}).$$

Related item: `_refln.A_meas_au` (conversion arbitrary).

[`refln`]

**`_refln.A_meas_au`** (float)

The measured value of structure-factor component  $A$  in arbitrary units.

$$A = |F| \cos(\text{phase}).$$

Related item: `_refln.A_meas` (conversion arbitrary).

[`refln`]

**`_refln.B_calc`** (float)  
`_refln_B_calc` (cif\_core.dic 2.0.1)  
 The calculated value of structure-factor component  $B$  in electrons.

$$B = |F| \sin(\text{phase}).$$

Related item: `_refln.B_calc_au` (conversion arbitrary). [refln]

**`_refln.B_calc_au`** (float)  
 The calculated value of structure-factor component  $B$  in arbitrary units.

$$B = |F| \sin(\text{phase}).$$

Related item: `_refln.B_calc` (conversion arbitrary). [refln]

**`_refln.B_meas`** (float)  
`_refln_B_meas` (cif\_core.dic 2.0.1)  
 The measured value of structure-factor component  $B$  in electrons.

$$B = |F| \sin(\text{phase}).$$

Related item: `_refln.B_meas_au` (conversion arbitrary). [refln]

**`_refln.B_meas_au`** (float)  
 The measured value of structure-factor component  $B$  in arbitrary units.

$$B = |F| \sin(\text{phase}).$$

Related item: `_refln.B_meas` (conversion arbitrary). [refln]

**`_refln.class_code`** (code)  
`_refln_class_code` (cif\_core.dic 2.3)  
 The code identifying the class to which this reflection has been assigned. This code must match a value of `_reflns.class_code`. Reflections may be grouped into classes for a variety of purposes. For example, for modulated structures each reflection class may be defined by the number  $m = \sum |m_i|$ , where the  $m_i$  are the integer coefficients that, in addition to  $h, k, l$ , index the corresponding diffraction vector in the basis defined for the reciprocal lattice.

[refln]

\* **`_refln.crystal_id`**  
`_refln_crystal_id` (cif\_core.dic 2.0.1)  
 This data item is a pointer to `_exptl_crystal_id` in the EXPTL\_CRYSTAL category.

**`_refln.d_spacing`** (float)  
`_refln_d_spacing` (cif\_core.dic 2.3)  
 The  $d$  spacing in ångströms for this reflection. This is related to the  $(\sin \theta)/\lambda$  value by the expression `_refln.d_spacing = 2/(_refln.sint/lambda)`.  
 The permitted range is  $[0.0, \infty)$ . [refln]

**`_refln.F_calc`** (float)  
`_refln_F_calc` (cif\_core.dic 2.0.1)  
 The calculated value of the structure factor in electrons.  
 Related item: `_refln.F_calc_au` (conversion arbitrary). [refln]

**`_refln.F_calc_au`** (float)  
 The calculated value of the structure factor in arbitrary units.  
 Related item: `_refln.F_calc` (conversion arbitrary). [refln]

**`_refln.F_meas`** (float, su)  
`_refln_F_meas` (cif\_core.dic 2.0.1)  
 The measured value of the structure factor in electrons.  
 Related items: `_refln.F_meas_sigma` (associated esd),  
`_refln.F_meas_au` (conversion arbitrary). [refln]

**`_refln.F_meas_au`** (float, su)  
 The measured value of the structure factor in arbitrary units.  
 Related items: `_refln.F_meas_sigma_au` (associated esd),  
`_refln.F_meas` (conversion arbitrary). [refln]

**`_refln.F_meas_sigma`** (float)  
`_refln_F_sigma` (cif\_core.dic 2.0.1)  
 The standard uncertainty (estimated standard deviation) of `_refln.F_meas` in electrons.  
 Related items: `_refln.F_meas` (associated value),  
`_refln.F_meas_sigma_au` (conversion arbitrary). [refln]

**`_refln.F_meas_sigma_au`** (float)  
 The standard uncertainty (estimated standard deviation) of `_refln.F_meas_au` in arbitrary units.  
 Related items: `_refln.F_meas_au` (associated value),  
`_refln.F_meas_sigma` (conversion arbitrary). [refln]

**`_refln.F_squared_calc`** (float)  
`_refln_F_squared_calc` (cif\_core.dic 2.0.1)  
 The calculated value of the squared structure factor in electrons squared.  
 [refln]

**`_refln.F_squared_meas`** (float)  
`_refln_F_squared_meas` (cif\_core.dic 2.0.1)  
 The measured value of the squared structure factor in electrons squared.  
 [refln]

**`_refln.F_squared_sigma`** (float)  
`_refln_F_squared_sigma` (cif\_core.dic 2.0.1)  
 The standard uncertainty (derived from measurement) of the squared structure factor in electrons squared.  
 [refln]

**`_refln.fom`** (float)  
 The figure of merit  $m$  for this reflection.

$$m = \frac{\int P_\alpha \exp(i\alpha) d\alpha}{\int P_\alpha d\alpha},$$

where  $P_\alpha$  = the probability that the phase angle  $\alpha$  is correct; the integral is taken over the range  $\alpha = 0$  to  $2\pi$ .

The permitted range is  $[0.0, \infty)$ . [refln]

**`_refln.include_status`** (code)  
`_refln_include_status` (cif\_core.dic 2.3)  
 Classification of a reflection so as to indicate its status with respect to inclusion in the refinement and the calculation of  $R$  factors.

Related item: `_refln.status` (alternate).

The data value must be one of the following:

- o (lower-case letter o for 'observed') satisfies `_refine.ls_d_res_high`, satisfies `_refine.ls_d_res_low` and exceeds `_reflns.threshold_expression`
- < satisfies `_refine.ls_d_res_high`, satisfies `_refine.ls_d_res_low` and does not exceed `_reflns.threshold_expression`
- systematically absent reflection
- x unreliable measurement – not used
- h does not satisfy `_refine.ls_d_res_high`
- l does not satisfy `_refine.ls_d_res_low`

[refln]

<p>* <b>_refln.index_h</b> (int)  <i>_refln_index_h</i> (cif_core.dic 2.0.1)                      Miller index <i>h</i> of the reflection. The values of the Miller indices in the REFLN category must correspond to the cell defined by cell lengths and cell angles in the CELL category.</p>	<p>* <b>_refln.scale_group_code</b> (float)  <i>_refln_scale_group_code</i> (cif_core.dic 2.0.1)                      This data item is a pointer to <i>_reflns_scale.group_code</i> in the REFLNS_SCALE category.</p>
<p>* <b>_refln.index_k</b> (int)  <i>_refln_index_k</i> (cif_core.dic 2.0.1)                      Miller index <i>k</i> of the reflection. The values of the Miller indices in the REFLN category must correspond to the cell defined by cell lengths and cell angles in the CELL category.</p>	<p><b>_refln.sint_over_lambda</b> (float)  <i>_refln_sint/lambda</i> (cif_core.dic 2.0.1)                      The <math>(\sin \theta)/\lambda</math> value in reciprocal ångströms for this reflection. The permitted range is [0.0, ∞).</p>
<p>* <b>_refln.index_l</b> (int)  <i>_refln_index_l</i> (cif_core.dic 2.0.1)                      Miller index <i>l</i> of the reflection. The values of the Miller indices in the REFLN category must correspond to the cell defined by cell lengths and cell angles in the CELL category.</p>	<p><b>_refln.status</b> (ucode)  <i>_refln_observed_status</i> (cif_core.dic 2.0.1)                      Classification of a reflection so as to indicate its status with respect to inclusion in the refinement and the calculation of <i>R</i> factors. The data value must be one of the following:</p> <ul style="list-style-type: none"> <li>o satisfies <i>_refine.ls_d_res_high</i>, satisfies <i>_refine.ls_d_res_low</i>, observed by <i>_reflns_observed_criterion</i>, not flagged as systematically absent, not flagged as unreliable</li> <li>&lt; satisfies <i>_refine.ls_d_res_high</i>, satisfies <i>_refine.ls_d_res_low</i>, unobserved by <i>_reflns_observed_criterion</i>, not flagged as systematically absent, not flagged as unreliable</li> <li>- systematically absent reflection</li> <li>x unreliable measurement – not used</li> <li>h does not satisfy <i>_refine.ls_d_res_high</i></li> <li>l does not satisfy <i>_refine.ls_d_res_low</i></li> <li>f satisfies <i>_refine.ls_d_res_high</i>, satisfies <i>_refine.ls_d_res_low</i>, observed by <i>_reflns_observed_criterion</i>, not flagged as systematically absent, not flagged as unreliable, excluded from refinement so as to be included in the calculation of a ‘free’ <i>R</i> factor</li> </ul>
<p><b>_refln.intensity_calc</b> (float)  <i>_refln_intensity_calc</i> (cif_core.dic 2.0.1)                      The calculated value of the intensity in the same units as <i>_refln.intensity_meas</i>.</p>	<p><b>_refln.symmetry_epsilon</b> (int)  <i>_refln_symmetry_epsilon</i> (cif_core.dic 2.0.1)                      The symmetry reinforcement factor corresponding to the number of times the reflection indices are generated identically from the space-group symmetry operations. The permitted range is [1, 48].</p>
<p><b>_refln.intensity_meas</b> (float)  <i>_refln_intensity_meas</i> (cif_core.dic 2.0.1)                      The measured value of the intensity.</p>	<p><b>_refln.symmetry_multiplicity</b> (int)  <i>_refln_symmetry_multiplicity</i> (cif_core.dic 2.0.1)                      The number of symmetry-equivalent reflections. The equivalent reflections have the same structure-factor magnitudes because of the space-group symmetry and the Friedel relationship. The permitted range is [1, 48].</p>
<p><b>_refln.intensity_sigma</b> (float)  <i>_refln_intensity_sigma</i> (cif_core.dic 2.0.1)                      The standard uncertainty (derived from measurement) of the intensity in the same units as <i>_refln.intensity_meas</i>.</p>	<p><b>_refln.wavelength</b> (float)  <i>_refln_wavelength</i> (cif_core.dic 2.0.1)                      The mean wavelength in ångströms of radiation used to measure this reflection. This is an important parameter for data collected using energy-dispersive detectors or the Laue method. The permitted range is [0.0, ∞).</p>
<p><b>_refln.mean_path_length_tbar</b> (float)  <i>_refln_mean_path_length_tbar</i> (cif_core.dic 2.3)                      Mean path length in millimetres through the crystal for this reflection. The permitted range is [0.0, ∞).</p>	<p>* <b>_refln.wavelength_id</b> (float)  <i>_refln_wavelength_id</i> (cif_core.dic 2.0.1)                      This data item is a pointer to <i>_diffrn_radiation.wavelength_id</i> in the DIFFRN_RADIATION category.</p>
<p><b>_refln.phase_calc</b> (float)  <i>_refln_phase_calc</i> (cif_core.dic 2.0.1)                      The calculated structure-factor phase in degrees.</p>	
<p><b>_refln.phase_meas</b> (float)  <i>_refln_phase_meas</i> (cif_core.dic 2.0.1)                      The measured structure-factor phase in degrees.</p>	
<p><b>_refln.refinement_status</b> (ucode)  <i>_refln_refinement_status</i> (cif_core.dic 2.0.1)                      Status of a reflection in the structure-refinement process. The data value must be one of the following:</p> <ul style="list-style-type: none"> <li>incl included in least-squares process</li> <li>excl excluded from least-squares process</li> <li>extn excluded due to extinction</li> </ul> <p>Where no value is given, the assumed value is ‘incl’.</p>	

**REFLN\_SYS\_ABS**

Data items in the REFLN\_SYS\_ABS category record details about the reflection data that should be systematically absent, given the designated space group.

Category group(s): `inclusive_group`  
`refln_group`

Category key(s): `_refln_sys_abs.index_h`  
`_refln_sys_abs.index_k`  
`_refln_sys_abs.index_l`

*Example 1 – hypothetical example.*

```
loop_
_refln_sys_abs.index_h
_refln_sys_abs.index_k
_refln_sys_abs.index_l
_refln_sys_abs.I
_refln_sys_abs.sigmaI
_refln_sys_abs.I_over_sigmaI
0 3 0 28.32 22.95 1.23
0 5 0 14.11 16.38 0.86
0 7 0 114.81 20.22 5.67
0 9 0 32.99 24.51 1.35
```

`_refln_sys_abs.I` (float, su)

`_ebi_refln_sys_abs.I` (*ebi\_extensions 1.0*)

The measured value of the intensity in arbitrary units.

Related item: `_refln_sys_abs.sigmaI` (associated esd). [refln\_sys\_abs]

`_refln_sys_abs.I_over_sigmaI` (float)

`_ebi_refln_sys_abs.I_over_sigmaI` (*ebi\_extensions 1.0*)

The ratio of `_refln_sys_abs.I` to `_refln_sys_abs.sigmaI`. Used to evaluate whether a reflection that should be systematically absent according to the designated space group is in fact absent.

[refln\_sys\_abs]

\* `_refln_sys_abs.index_h` (int)

`_ebi_refln_sys_abs.h` (*ebi\_extensions 1.0*)

Miller index *h* of the reflection. The values of the Miller indices in the REFLN\_SYS\_ABS category must correspond to the cell defined by cell lengths and cell angles in the CELL category.

[refln\_sys\_abs]

\* `_refln_sys_abs.index_k` (int)

`_ebi_refln_sys_abs.k` (*ebi\_extensions 1.0*)

Miller index *k* of the reflection. The values of the Miller indices in the REFLN\_SYS\_ABS category must correspond to the cell defined by cell lengths and cell angles in the CELL category.

[refln\_sys\_abs]

\* `_refln_sys_abs.index_l` (int)

`_ebi_refln_sys_abs.l` (*ebi\_extensions 1.0*)

Miller index *l* of the reflection. The values of the Miller indices in the REFLN\_SYS\_ABS category must correspond to the cell defined by cell lengths and cell angles in the CELL category.

[refln\_sys\_abs]

`_refln_sys_abs.sigmaI` (float)

`_ebi_refln_sys_abs.sigmaI` (*ebi\_extensions 1.0*)

The standard uncertainty (estimated standard deviation) of `_refln_sys_abs.I` in arbitrary units.

Related item: `_refln_sys_abs.I` (associated value). [refln\_sys\_abs]

**REFLNS**

Data items in the REFLNS category record details about the reflection data used to determine the ATOM\_SITE data items. The REFLNS data items refer to individual reflections and must be included in looped lists. The REFLNS data items specify the parameters that apply to all reflections. The REFLNS data items are not looped.

Category group(s): `inclusive_group`  
`refln_group`

Category key(s): `_reflns.entry_id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
_reflns.entry_id          '5HVP'
_reflns.data_reduction_method
; Xengen program scalei. Anomalous pairs were merged. Scaling
proceeded in several passes, beginning with 1-parameter
fit and ending with 3-parameter fit.
;
_reflns.data_reduction_details
; Merging and scaling based on only those reflections
with I > \s(I).
;

_reflns.d_resolution_high      2.00
_reflns.d_resolution_low      8.00

_reflns.limit_h_max           22
_reflns.limit_h_min           0
_reflns.limit_k_max           46
_reflns.limit_k_min           0
_reflns.limit_l_max           57
_reflns.limit_l_min           0

_reflns.number_obs            7228
_reflns.observed_criterion    '> 1 \s(I)'
_reflns.details                none
```

*Example 2 – based on data set TOZ of Willis, Beckwith & Tozer [Acta Cryst. (1991), C47, 2276–2277].*

```
_reflns.limit_h_min         0
_reflns.limit_h_max         6
_reflns.limit_k_min         0
_reflns.limit_k_max         17
_reflns.limit_l_min         0
_reflns.limit_l_max         22
_reflns.number_all          1592
_reflns.number_obs          1408
_reflns.observed_criterion   'F > 6.0_\s(F)'
_reflns.d_resolution_high    0.8733
_reflns.d_resolution_low     11.9202
```

`_reflns.B_iso_wilson_estimate` (float)

The value of the overall isotropic displacement parameter estimated from the slope of the Wilson plot.

[reflns]

`_reflns.d_resolution_high` (float)

`_reflns_d_resolution_high` (*cif\_core.dic 2.0.1*)

The smallest value for the interplanar spacings for the reflection data. This is called the highest resolution.

The permitted range is [0.0, ∞).

[reflns]

`_reflns.d_resolution_low` (float)

`_reflns_d_resolution_low` (*cif\_core.dic 2.0.1*)

The largest value for the interplanar spacings for the reflection data. This is called the lowest resolution.

The permitted range is [0.0, ∞).

[reflns]

`_reflns.data_reduction_details` (text)

A description of special aspects of the data-reduction procedures.

Example:

```
; Merging and scaling based on only those
reflections with I > sig(I).
```

[reflns]

**`_reflns.data_reduction_method`** (text)  
 The method used for data reduction. Note that this is not the computer program used, which is described in the SOFTWARE category, but the method itself. This data item should be used to describe significant methodological options used within the data-reduction programs.

Example:

```
; Profile fitting by method of Kabsch (1987).
; Scaling used spherical harmonic coefficients.
; [reflns]
```

**`_reflns.details`** (text)  
**`_reflns_special_details`** (cif\_core.dic 2.0.1)  
 A description of reflection data not covered by other data names. This should include details of the Friedel pairs.

[reflns]

\* **`_reflns.entry_id`**  
 This data item is a pointer to `_entry.id` in the ENTRY category.

**`_reflns.Friedel_coverage`** (float)  
**`_reflns_Friedel_coverage`** (cif\_core.dic 2.3)  
 The proportion of Friedel-related reflections present in the number of 'independent' reflections specified by the item `_reflns.number_all`. This proportion is calculated as the ratio

$$\frac{[N(\text{crystal class}) - N(\text{Laue symmetry})]}{N(\text{Laue symmetry})},$$

where, working from the DIFFRN\_REFLN list,  $N(\text{Crystal class})$  is the number of reflections obtained on averaging under the symmetry of the crystal class and  $N(\text{Laue symmetry})$  is the number of reflections obtained on averaging under the Laue symmetry.

Examples: (a) For centrosymmetric structures, the value of `_reflns.Friedel_coverage` is necessarily equal to 0.0 as the crystal class is identical to the Laue symmetry. (b) For whole-sphere data for a crystal in the space group  $P1$ , `_reflns.Friedel_coverage` is equal to 1.0, as no reflection  $hkl$  is equivalent to  $-h - k - l$  in the crystal class and all Friedel pairs  $\{hkl; -h - k - l\}$  have been measured. (c) For whole-sphere data in space group  $Pmm2$ , `_reflns.Friedel_coverage` will be  $< 1.0$  because although reflections  $hkl$  and  $-h - k - l$  are not equivalent when  $hkl$  indices are nonzero, they are when  $l = 0$ . (d) For a crystal in space group  $Pmm2$ , measurements of the two inequivalent octants  $h \geq 0, k \geq 0, l$  lead to the same value as in (c), whereas measurements of the two equivalent octants  $h \geq 0, k, l \geq 0$  will lead to a value of zero for `_reflns.Friedel_coverage`.

The permitted range is [0.0, 1.0]. [reflns]

**`_reflns.limit_h_max`** (int)  
**`_reflns_limit_h_max`** (cif\_core.dic 2.0.1)  
 Maximum value of the Miller index  $h$  for the reflection data. This need not have the same value as `_diffrn_reflns.limit_h_max`.

[reflns]

**`_reflns.limit_h_min`** (int)  
**`_reflns_limit_h_min`** (cif\_core.dic 2.0.1)  
 Minimum value of the Miller index  $h$  for the reflection data. This need not have the same value as `_diffrn_reflns.limit_h_min`.

[reflns]

**`_reflns.limit_k_max`** (int)  
**`_reflns_limit_k_max`** (cif\_core.dic 2.0.1)  
 Maximum value of the Miller index  $k$  for the reflection data. This need not have the same value as `_diffrn_reflns.limit_k_max`.

[reflns]

**`_reflns.limit_k_min`** (int)  
**`_reflns_limit_k_min`** (cif\_core.dic 2.0.1)  
 Minimum value of the Miller index  $k$  for the reflection data. This need not have the same value as `_diffrn_reflns.limit_k_min`.

[reflns]

**`_reflns.limit_l_max`** (int)  
**`_reflns_limit_l_max`** (cif\_core.dic 2.0.1)  
 Maximum value of the Miller index  $l$  for the reflection data. This need not have the same value as `_diffrn_reflns.limit_l_max`.

[reflns]

**`_reflns.limit_l_min`** (int)  
**`_reflns_limit_l_min`** (cif\_core.dic 2.0.1)  
 Minimum value of the Miller index  $l$  for the reflection data. This need not have the same value as `_diffrn_reflns.limit_l_min`.

[reflns]

**`_reflns.number_all`** (int)  
**`_reflns_number_total`** (cif\_core.dic 2.0.1)  
 The total number of reflections in the REFLN list (not the DIFFRN\_REFLN list). This number may contain Friedel-equivalent reflections according to the nature of the structure and the procedures used. The item `_reflns.details` describes the reflection data.

The permitted range is  $[0, \infty)$ . [reflns]

**`_reflns.number_gt`** (int)  
**`_reflns_number_gt`** (cif\_core.dic 2.3)  
 The number of reflections in the REFLN list (not the DIFFRN\_REFLN list) that are significantly intense, satisfying the criterion specified by `_reflns.threshold_expression`. This may include Friedel-equivalent reflections (*i.e.* those which are symmetry-equivalent under the Laue symmetry but inequivalent under the crystal class) according to the nature of the structure and the procedures used. Any special characteristics of the reflections included in the REFLN list should be described using the item `_reflns.details`.

The permitted range is  $[0, \infty)$ . [reflns]

**`_reflns.number_obs`** (int)  
**`_reflns_number_observed`** (cif\_core.dic 2.0.1)  
 The number of reflections in the REFLN list (not the DIFFRN\_REFLN list) classified as observed (see `_reflns.observed_criterion`). This number may contain Friedel-equivalent reflections according to the nature of the structure and the procedures used.

The permitted range is  $[0, \infty)$ . [reflns]

**`_reflns.observed_criterion`** (text)  
**`_reflns_observed_criterion`** (cif\_core.dic 2.0.1)  
 The criterion used to classify a reflection as 'observed'. This criterion is usually expressed in terms of a  $\sigma(I)$  or  $\sigma(F)$  threshold.

Related items: `_reflns.observed_criterion_sigma_F` (alternate),

`_reflns.observed_criterion_sigma_I` (alternate),

`_reflns.observed_criterion_I_min` (alternate),

`_reflns.observed_criterion_I_max` (alternate),

`_reflns.observed_criterion_F_min` (alternate),

`_reflns.observed_criterion_F_max` (alternate).

Example: '>2sigma (I)'. [reflns]

**`_reflns.observed_criterion_F_max`** (float)  
 The criterion used to classify a reflection as 'observed' expressed as an upper limit for the value of  $F$ .

Related items: `_reflns.observed_criterion` (alternate),

`_reflns.observed_criterion_I_max` (convention). [reflns]

**`_reflns.observed_criterion_F_min`** (float)  
 The criterion used to classify a reflection as ‘observed’ expressed as a lower limit for the value of  $F$ .  
 Related items: `_reflns.observed_criterion` (alternate),  
`_reflns.observed_criterion_I_min` (convention). [reflns]

**`_reflns.observed_criterion_I_max`** (float)  
 The criterion used to classify a reflection as ‘observed’ expressed as an upper limit for the value of  $I$ .  
 Related items: `_reflns.observed_criterion` (alternate),  
`_reflns.observed_criterion_F_max` (convention). [reflns]

**`_reflns.observed_criterion_I_min`** (float)  
 The criterion used to classify a reflection as ‘observed’ expressed as a lower limit for the value of  $I$ .  
 Related items: `_reflns.observed_criterion` (alternate),  
`_reflns.observed_criterion_F_min` (convention). [reflns]

**`_reflns.observed_criterion_sigma_F`** (float)  
 The criterion used to classify a reflection as ‘observed’ expressed as a multiple of the value of  $\sigma(F)$ .  
 Related items: `_reflns.observed_criterion` (alternate),  
`_reflns.observed_criterion_sigma_I` (convention). [reflns]

**`_reflns.observed_criterion_sigma_I`** (float)  
 The criterion used to classify a reflection as ‘observed’ expressed as a multiple of the value of  $\sigma(I)$ .  
 Related items: `_reflns.observed_criterion` (alternate),  
`_reflns.observed_criterion_sigma_F` (convention). [reflns]

**`_reflns.percent_possible_obs`** (float)  
 The percentage of geometrically possible reflections represented by reflections that satisfy the resolution limits established by `_reflns.d_resolution_high` and `_reflns.d_resolution_low` and the observation limit established by `_reflns.observed_criterion`.  
 The permitted range is  $[0.0, \infty)$ . [reflns]

**`_reflns.R_free_details`** (text)  
 A description of the method by which a subset of reflections was selected for exclusion from refinement so as to be used in the calculation of a ‘free’  $R$  factor.  
 Example:  
 ; The data set was sorted with l varying most rapidly and h  
 ; varying least rapidly. Every 10th reflection in this sorted  
 ; list was excluded from refinement and included in the  
 ; calculation of a ‘free’ R factor.  
 ; [reflns]

**`_reflns.Rmerge_F_all`** (float)  
 Residual factor  $R_{\text{merge}}$  for all reflections that satisfy the resolution limits established by `_reflns.d_resolution_high` and `_reflns.d_resolution_low`.

$$R_{\text{merge}} = \frac{\sum_i (\sum_j |F_j - \langle F \rangle|)}{\sum_i (\sum_j \langle F \rangle)}$$

where  $F_j$  = the amplitude of the  $j$ th observation of reflection  $i$ ,  $\langle F \rangle$  = the mean of the amplitudes of all observations of reflection  $i$ ,  $\sum_i$  is taken over all reflections and  $\sum_j$  is taken over all observations of each reflection.  
 The permitted range is  $[0.0, \infty)$ . [reflns]

**`_reflns.Rmerge_F_obs`** (float)  
 Residual factor  $R_{\text{merge}}$  for reflections that satisfy the resolution limits established by `_reflns.d_resolution_high` and `_reflns.d_resolution_low` and the observation limit established by `_reflns.observed_criterion`.

$$R_{\text{merge}} = \frac{\sum_i (\sum_j |F_j - \langle F \rangle|)}{\sum_i (\sum_j \langle F \rangle)}$$

where  $F_j$  = the amplitude of the  $j$ th observation of reflection  $i$ ,  $\langle F \rangle$  = the mean of the amplitudes of all observations of reflection  $i$ ,  $\sum_i$  is taken over all reflections and  $\sum_j$  is taken over all observations of each reflection.  
 The permitted range is  $[0.0, \infty)$ . [reflns]

**`_reflns.threshold_expression`** (text)  
**`_reflns.threshold_expression`** (cif\_core.dic 2.3)  
 The threshold, usually based on multiples of  $u(I)$ ,  $u(F^2)$  or  $u(F)$ , that serves to identify significantly intense reflections, the number of which is given by `_reflns.number_gt`. These reflections are used in the calculation of `_refine.ls_R_factor_gt`.  
 Related item: `_reflns.observed_criterion` (alternate).  
 Example: ‘I>2u (I)’. [reflns]

### REFLNS\_CLASS

Data items in the REFLNS\_CLASS category record details of the reflections used to determine the structural parameters for each reflection class.

Category key(s): `_reflns_class.code`

*Example 1 – example corresponding to the one-dimensional incommensurately modulated structure of  $K_2\text{SeO}_4$ .*

```
loop_
  _reflns_class.number_gt
  _reflns_class.code
    584 'Main'
    226 'Sat1'
    50  'Sat2'
```

\* **`_reflns_class.code`** (code)  
**`_reflns_class.code`** (cif\_core.dic 2.3)  
 The code identifying a certain reflection class.  
 Examples: ‘1’, ‘m1’, ‘s2’. [reflns\_class]

**`_reflns_class.d_res_high`** (float)  
**`_reflns_class.d_res_high`** (cif\_core.dic 2.3)  
 For each reflection class, the smallest value in ångströms for the interplanar spacings for the reflections used in the refinement. This is called the highest resolution.  
 The permitted range is  $[0.0, \infty)$ . [reflns\_class]

**`_reflns_class.d_res_low`** (float)  
**`_reflns_class.d_res_low`** (cif\_core.dic 2.3)  
 For each reflection class, the largest value in ångströms for the interplanar spacings for the reflections used in the refinement. This is called the lowest resolution.  
 The permitted range is  $[0.0, \infty)$ . [reflns\_class]

**`_reflns_class.description`** (text)  
**`_reflns_class.description`** (cif\_core.dic 2.3)  
 Description of each reflection class.  
 Examples: ‘m=1 first order satellites’,  
 ‘H0L0 common projection reflections’. [reflns\_class]

**\_reflns\_class.number\_gt** (int)  
**\_reflns\_class\_number\_gt** (cif\_core.dic 2.3)  
 For each reflection class, the number of significantly intense reflections (see **\_reflns.threshold\_expression**) in the REFLN list (not the DIFFRN\_REFLN list). This may include Friedel-equivalent reflections (*i.e.* those which are symmetry-equivalent under the Laue symmetry but inequivalent under the crystal class) according to the nature of the structure and the procedures used. Any special characteristics of the reflections included in the REFLN list should be described using the item **\_reflns.details**.  
 The permitted range is [0, ∞). [reflns\_class]

**\_reflns\_class.number\_total** (int)  
**\_reflns\_class\_number\_total** (cif\_core.dic 2.3)  
 For each reflection class, the total number of reflections in the REFLN list (not the DIFFRN\_REFLN list). This may include Friedel-equivalent reflections (*i.e.* those which are symmetry-equivalent under the Laue symmetry but inequivalent under the crystal class) according to the nature of the structure and the procedures used. Any special characteristics of the reflections included in the REFLN list should be described using the item **\_reflns.details**.  
 The permitted range is [0, ∞). [reflns\_class]

**\_reflns\_class.R\_factor\_all** (float)  
**\_reflns\_class\_R\_factor\_all** (cif\_core.dic 2.3)  
 For each reflection class, the residual factor for all reflections included in the refinement. The reflections also satisfy the resolution limits established by **\_reflns\_class.d\_res\_high** and **\_reflns\_class.d\_res\_low**. This is the conventional *R* factor. See also the definition of **\_reflns\_class.wR\_factor\_all**.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|}$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the reflections of this class.  
 The permitted range is [0.0, ∞). [reflns\_class]

**\_reflns\_class.R\_factor\_gt** (float)  
**\_reflns\_class\_R\_factor\_gt** (cif\_core.dic 2.3)  
 For each reflection class, the residual factor for significantly intense reflections (see **\_reflns.threshold\_expression**) included in the refinement. The reflections also satisfy the resolution limits established by **\_reflns\_class.d\_res\_high** and **\_reflns\_class.d\_res\_low**. This is the conventional *R* factor. See also the definition of **\_reflns\_class.wR\_factor\_all**.

$$R = \frac{\sum |F_{\text{obs}} - F_{\text{calc}}|}{\sum |F_{\text{obs}}|}$$

where  $F_{\text{obs}}$  = the observed structure-factor amplitudes,  $F_{\text{calc}}$  = the calculated structure-factor amplitudes and the sum is taken over the reflections of this class.  
 The permitted range is [0.0, ∞). [reflns\_class]

**\_reflns\_class.R\_Fsqd\_factor** (float)  
**\_reflns\_class\_R\_Fsqd\_factor** (cif\_core.dic 2.3)  
 For each reflection class, the residual factor  $R(F^2)$  calculated on the squared amplitudes of the observed and calculated structure factors for the reflections judged significantly intense (*i.e.* satisfying the threshold specified by **\_reflns.threshold\_expression**) and included in the refinement. The reflections also satisfy the

resolution limits established by **\_reflns\_class.d\_res\_high** and **\_reflns\_class.d\_res\_low**.

$$R(F^2) = \frac{\sum |F_{\text{obs}}^2 - F_{\text{calc}}^2|}{\sum |F_{\text{obs}}^2|}$$

where  $F_{\text{obs}}^2$  = squares of the observed structure-factor amplitudes,  $F_{\text{calc}}^2$  = squares of the calculated structure-factor amplitudes and the sum is taken over the reflections of this class.  
 The permitted range is [0.0, ∞). [reflns\_class]

**\_reflns\_class.R\_I\_factor** (float)  
**\_reflns\_class\_R\_I\_factor** (cif\_core.dic 2.3)  
 For each reflection class, the residual factor  $R(I)$  for the reflections judged significantly intense (*i.e.* satisfying the threshold specified by **\_reflns.threshold\_expression**) and included in the refinement. This is most often calculated in Rietveld refinements against powder data, where it is referred to as  $R_B$  or  $R_{\text{Bragg}}$ .

$$R(I) = \frac{\sum |I_{\text{obs}} - I_{\text{calc}}|}{\sum |I_{\text{obs}}|}$$

where  $I_{\text{obs}}$  = the net observed intensities,  $I_{\text{calc}}$  = the net calculated intensities and the sum is taken over the reflections of this class.  
 The permitted range is [0.0, ∞). [reflns\_class]

**\_reflns\_class.wR\_factor\_all** (float)  
**\_reflns\_class\_wR\_factor\_all** (cif\_core.dic 2.3)  
 For each reflection class, the weighted residual factors for all reflections included in the refinement. The reflections also satisfy the resolution limits established by **\_reflns\_class.d\_res\_high** and **\_reflns\_class.d\_res\_low**. See also **\_reflns\_class.R\_factor** definitions.

$$wR = \left( \frac{\sum |w|Y_{\text{obs}} - Y_{\text{calc}}|^2|}{\sum |wY_{\text{obs}}^2|} \right)^{1/2}$$

where  $Y_{\text{obs}}$  = the observed amplitude specified by **\_refine.ls\_structure\_factor\_coef**,  $Y_{\text{calc}}$  = the calculated amplitude specified by **\_refine.ls\_structure\_factor\_coef**,  $w$  = the least-squares weight and the sum is taken over the reflections of this class.  
 The permitted range is [0.0, ∞). [reflns\_class]

REFLNS_SCALE	
Data items in the REFLNS_SCALE category record details about the structure-factor scales. They are referenced from within the REFLN list through <b>_refln.scale_group_code</b> .	
Category group(s): <b>inclusive_group</b> <b>refln_group</b>	
Category key(s): <b>_reflns_scale.group_code</b>	
<i>Example 1 – based on laboratory records for the collagen-like peptide [(POG)<sub>n</sub>EKG (POG)<sub>5</sub>]<sub>3</sub>.</i>	
<b>_reflns_scale.group_code</b>	SG1
<b>_reflns_scale.meas_F</b>	4.0

\* **\_reflns\_scale.group\_code** (line)  
**\_reflns\_scale\_group\_code** (cif\_core.dic 2.0.1)  
 The code identifying a scale **\_reflns\_scale.meas\_F**, **\_reflns\_scale.meas\_F\_squared** OR **\_reflns\_scale.meas\_intensity**. These are linked to the REFLN list by the **\_refln.scale\_group\_code**. These codes need not correspond to those in the DIFFRN\_SCALE list.  
 The following item(s) have an equivalent role in their respective categories:  
**\_refln.scale\_group\_code**.  
 Examples: '1', '2', 'c1', 'c2'. [reflns\_scale]

**reflns\_scale.meas\_F** (float)  
reflns\_scale\_meas\_F (cif.core.dic 2.0.1)

A scale associated with reflns\_scale.group\_code.  
 The permitted range is [0.0, ∞). [reflns\_scale]

**reflns\_scale.meas\_F\_squared** (float)  
reflns\_scale\_meas\_F\_squared (cif.core.dic 2.0.1)

A scale associated with reflns\_scale.group\_code.  
 The permitted range is [0.0, ∞). [reflns\_scale]

**reflns\_scale.meas\_intensity** (float)  
reflns\_scale\_meas\_intensity (cif.core.dic 2.0.1)

A scale associated with reflns\_scale.group\_code.  
 The permitted range is [0.0, ∞). [reflns\_scale]

### REFLNS\_SHELL

Data items in the REFLNS\_SHELL category record details about the reflection data used to determine the ATOM\_SITE data items broken down into shells of resolution.

Category group(s): inclusive\_group  
refln\_group

Category key(s): reflns\_shell.d\_res\_high  
reflns\_shell.d\_res\_low

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  reflns_shell.d_res_high
  reflns_shell.d_res_low
  reflns_shell.meanI_over_sigI_obs
  reflns_shell.number_measured_obs
  reflns_shell.number_unique_obs
  reflns_shell.percent_possible_obs
  reflns_shell.Rmerge_F_obs
  31.38 3.82 69.8 9024 2540 96.8 1.98
  3.82 3.03 26.1 7413 2364 95.1 3.85
  3.03 2.65 10.5 5640 2123 86.2 6.37
  2.65 2.41 6.4 4322 1882 76.8 8.01
  2.41 2.23 4.3 3247 1714 70.4 9.86
  2.23 2.10 3.1 1140 812 33.3 13.99
```

\* **reflns\_shell.d\_res\_high** (float)  
reflns\_shell\_d\_res\_high (cif.core.dic 2.0.1)

The smallest value in ångströms for the interplanar spacings for the reflections in this shell. This is called the highest resolution.  
 The permitted range is [0.0, ∞). [reflns\_shell]

\* **reflns\_shell.d\_res\_low** (float)  
reflns\_shell\_d\_res\_low (cif.core.dic 2.0.1)

The highest value in ångströms for the interplanar spacings for the reflections in this shell. This is called the lowest resolution.  
 The permitted range is [0.0, ∞). [reflns\_shell]

**reflns\_shell.meanI\_over\_sigI\_all** (float)  
reflns\_shell\_meanI\_over\_sigI\_all (cif.core.dic 2.0.1)

The ratio of the mean of the intensities of all reflections in this shell to the mean of the standard uncertainties of the intensities of all reflections in this shell. [reflns\_shell]

**reflns\_shell.meanI\_over\_sigI\_gt** (float)  
reflns\_shell\_meanI\_over\_sigI\_gt (cif.core.dic 2.3)

The ratio of the mean of the intensities of the significantly intense reflections (see reflns.threshold\_expression) in this shell to the mean of the standard uncertainties of the intensities of the significantly intense reflections in this shell.

Related item: reflns\_shell.meanI\_over\_uI\_gt (replaces). [reflns\_shell]

**reflns\_shell.meanI\_over\_sigI\_obs** (float)  
reflns\_shell\_meanI\_over\_sigI\_obs (cif.core.dic 2.0.1)

The ratio of the mean of the intensities of the reflections classified as ‘observed’ (see reflns.observed\_criterion) in this shell to the mean of the standard uncertainties of the intensities of the ‘observed’ reflections in this shell.

**reflns\_shell.meanI\_over\_uI\_all** (float)  
reflns\_shell\_meanI\_over\_uI\_all (cif.core.dic 2.3)

The ratio of the mean of the intensities of all reflections in this shell to the mean of the standard uncertainties of the intensities of all reflections in this shell.

Related item: reflns\_shell.meanI\_over\_sigI\_all (alternate).

[reflns\_shell]

**reflns\_shell.meanI\_over\_uI\_gt** (float)  
reflns\_shell\_meanI\_over\_uI\_gt (cif.core.dic 2.3)

The ratio of the mean of the intensities of the significantly intense reflections (see reflns.threshold\_expression) in this shell to the mean of the standard uncertainties of the intensities of the significantly intense reflections in this shell.

Related items: reflns\_shell.meanI\_over\_sigI\_gt (alternate),

reflns\_shell.meanI\_over\_sigI\_obs (alternate). [reflns\_shell]

**reflns\_shell.number\_measured\_all** (int)  
reflns\_shell\_number\_measured\_all (cif.core.dic 2.0.1)

The total number of reflections measured for this shell.

[reflns\_shell]

**reflns\_shell.number\_measured\_gt** (int)  
reflns\_shell\_number\_measured\_gt (cif.core.dic 2.3)

The number of significantly intense reflections (see reflns.threshold\_expression) measured for this shell.

The permitted range is [0, ∞).

Related item: reflns\_shell.number\_measured\_obs (alternate).

[reflns\_shell]

**reflns\_shell.number\_measured\_obs** (int)  
reflns\_shell\_number\_measured\_obs (cif.core.dic 2.0.1)

The number of reflections classified as ‘observed’ (see reflns.observed\_criterion) for this shell.

[reflns\_shell]

**reflns\_shell.number\_possible** (int)  
reflns\_shell\_number\_possible (cif.core.dic 2.0.1)

The number of unique reflections it is possible to measure in this shell.

The permitted range is [0, ∞).

[reflns\_shell]

**reflns\_shell.number\_unique\_all** (int)  
reflns\_shell\_number\_unique\_all (cif.core.dic 2.0.1)

The total number of measured reflections which are symmetry-unique after merging for this shell.

[reflns\_shell]

**reflns\_shell.number\_unique\_gt** (int)  
reflns\_shell\_number\_unique\_gt (cif.core.dic 2.3)

The total number of significantly intense reflections (see reflns.threshold\_expression) resulting from merging measured symmetry-equivalent reflections for this resolution shell.

The permitted range is [0, ∞).

Related item: reflns\_shell.number\_unique\_obs (alternate). [reflns\_shell]

**`_reflns_shell.number_unique_obs`** (int)  
`_reflns_shell_number_unique_obs` (cif\_core.dic 2.0.1)

The total number of measured reflections classified as ‘observed’ (see `_reflns.observed_criterion`) which are symmetry-unique after merging for this shell.

[reflns\_shell]

**`_reflns_shell.percent_possible_all`** (float)  
`_reflns_shell_percent_possible_all` (cif\_core.dic 2.0.1)

The percentage of geometrically possible reflections represented by all reflections measured for this shell.

The permitted range is [0.0, ∞).

[reflns\_shell]

**`_reflns_shell.percent_possible_gt`** (float)  
`_reflns_shell_percent_possible_gt` (cif\_core.dic 2.3)

The percentage of geometrically possible reflections represented by significantly intense reflections (see `_reflns.threshold_expression`) measured for this shell.

The permitted range is [0.0, 100.0].

Related item: `_reflns_shell.percent_possible_obs` (alternate).

[reflns\_shell]

**`_reflns_shell.percent_possible_obs`** (float)  
`_reflns_shell_percent_possible_obs` (cif\_core.dic 2.0.1)

The percentage of geometrically possible reflections represented by reflections classified as ‘observed’ (see `_reflns.observed_criterion`) for this shell.

The permitted range is [0.0, ∞).

[reflns\_shell]

**`_reflns_shell.Rmerge_F_all`** (float)  
`_reflns_shell_Rmerge_F_all` (cif\_core.dic 2.0.1)

Residual factor  $R_{\text{merge}}$  for all reflections that satisfy the resolution limits established by `_reflns_shell.d_res_high` and `_reflns_shell.d_res_low`.

$$R_{\text{merge}} = \frac{\sum_i (\sum_j |F_j - \langle F \rangle|)}{\sum_i (\sum_j \langle F \rangle)}$$

where  $F_j$  = the amplitude of the  $j$ th observation of reflection  $i$ ,  $\langle F \rangle$  = the mean of the amplitudes of all observations of reflection  $i$ ,  $\sum_i$  is taken over all reflections and  $\sum_j$  is taken over all observations of each reflection.

The permitted range is [0.0, ∞).

[reflns\_shell]

**`_reflns_shell.Rmerge_F_gt`** (float)  
`_reflns_shell_Rmerge_F_gt` (cif\_core.dic 2.3)

The value of  $R_{\text{merge}}(F)$  for significantly intense reflections (see `_reflns.threshold_expression`) in a given shell.

$$R_{\text{merge}} = \frac{\sum_i (\sum_j |F_j - \langle F \rangle|)}{\sum_i (\sum_j \langle F \rangle)}$$

where  $F_j$  = the amplitude of the  $j$ th observation of reflection  $i$ ,  $\langle F \rangle$  = the mean of the amplitudes of all observations of reflection  $i$ ,  $\sum_i$  is taken over all reflections and  $\sum_j$  is taken over all observations of each reflection.

The permitted range is [0.0, ∞).

Related item: `_reflns_shell.Rmerge_F_obs` (alternate).

[reflns\_shell]

**`_reflns_shell.Rmerge_F_obs`** (float)  
`_reflns_shell_Rmerge_F_obs` (cif\_core.dic 2.0.1)

Residual factor  $R_{\text{merge}}$  for reflections that satisfy the resolution limits established by `_reflns_shell.d_res_high` and `_reflns_shell.d_res_low` and the observation criterion established by `_reflns.observed_criterion`.

$$R_{\text{merge}} = \frac{\sum_i (\sum_j |F_j - \langle F \rangle|)}{\sum_i (\sum_j \langle F \rangle)}$$

where  $F_j$  = the amplitude of the  $j$ th observation of reflection  $i$ ,  $\langle F \rangle$  = the mean of the amplitudes of all observations of reflection  $i$ ,  $\sum_i$  is taken over all reflections and  $\sum_j$  is taken over all observations of each reflection.

The permitted range is [0.0, ∞).

[reflns\_shell]

**`_reflns_shell.Rmerge_I_all`** (float)  
`_reflns_shell_Rmerge_I_all` (cif\_core.dic 2.0.1)

The value of  $R_{\text{merge}}(I)$  for all reflections in a given shell.

$$R_{\text{merge}}(I) = \frac{\sum_i (\sum_j |I_j - \langle I \rangle|)}{\sum_i (\sum_j \langle I \rangle)}$$

where  $I_j$  = the intensity of the  $j$ th observation of reflection  $i$ ,  $\langle I \rangle$  = the mean of the intensities of all observations of reflection  $i$ ,  $\sum_i$  is taken over all reflections and  $\sum_j$  is taken over all observations of each reflection.

The permitted range is [0.0, ∞).

[reflns\_shell]

**`_reflns_shell.Rmerge_I_gt`** (float)  
`_reflns_shell_Rmerge_I_gt` (cif\_core.dic 2.3)

The value of  $R_{\text{merge}}(I)$  for significantly intense reflections (see `_reflns.threshold_expression`) in a given shell.

$$R_{\text{merge}}(I) = \frac{\sum_i (\sum_j |I_j - \langle I \rangle|)}{\sum_i (\sum_j \langle I \rangle)}$$

where  $I_j$  = the intensity of the  $j$ th observation of reflection  $i$ ,  $\langle I \rangle$  = the mean of the intensities of all observations of reflection  $i$ ,  $\sum_i$  is taken over all reflections and  $\sum_j$  is taken over all observations of each reflection.

The permitted range is [0.0, ∞).

Related item: `_reflns_shell.Rmerge_I_obs` (alternate).

[reflns\_shell]

**`_reflns_shell.Rmerge_I_obs`** (float)  
`_reflns_shell_Rmerge_I_obs` (cif\_core.dic 2.0.1)

The value of  $R_{\text{merge}}(I)$  for reflections classified as ‘observed’ (see `_reflns.observed_criterion`) in a given shell.

$$R_{\text{merge}}(I) = \frac{\sum_i (\sum_j |I_j - \langle I \rangle|)}{\sum_i (\sum_j \langle I \rangle)}$$

where  $I_j$  = the intensity of the  $j$ th observation of reflection  $i$ ,  $\langle I \rangle$  = the mean of the intensities of all observations of reflection  $i$ ,  $\sum_i$  is taken over all reflections and  $\sum_j$  is taken over all observations of each reflection.

The permitted range is [0.0, ∞).

[reflns\_shell]

**SOFTWARE**

Data items in the SOFTWARE category record details about the software used in the structure analysis, which implies any software used in the generation of any data items associated with the structure determination and structure representation. These data items allow computer programs to be referenced in more detail than data items in the COMPUTING category do.

Category group(s): `inclusive_group`  
                   `computing_group`  
 Category key(s): `_software.name`  
                   `_software.version`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _software.name
  _software.version
  _software.date
  _software.type
  _software.contact_author
  _software.contact_author_email
  _software.location
  _software.classification
  _software.citation_id
  _software.language
  _software.compiler_name
  _software.compiler_version
  _software.hardware
  _software.os
  _software.os_version
  _software.dependencies
  _software.mods
  _software.description
  Prolsq unknown . program 'Wayne A. Hendrickson' ?
  'ftp://rosebud.sdsc.edu/pub/sdsc/xtal/CCP4/ccp4/'
  refinement ref5 Fortran
  'Convex Fortran' v8.0 'Convex C220' ConvexOS v10.1
  'Requires that Protin be run first' optimized
  'restrained least-squares refinement'
```

**\_software.citation\_id**

This data item is a pointer to `_citation.id` in the CITATION category.

**\_software.classification** (uline)

The classification of the program according to its major function.

Examples: 'data collection', 'data reduction', 'phasing',  
 'model building', 'refinement', 'validation', 'other'. [software]

**\_software.compiler\_name** (line)

The compiler used to compile the software.

Examples: 'Convex Fortran', 'gcc', 'DEC C'. [software]

**\_software.compiler\_version** (line)

The version of the compiler used to compile the software.

Examples: '3.1', '2.1 alpha'. [software]

**\_software.contact\_author** (line)

The recognized contact author of the software. This could be the original author, someone who has modified the code or someone who maintains the code. It should be the person most commonly associated with the code.

Examples: 'T. Alwyn Jones', 'Axel Brunger'. [software]

**\_software.contact\_author\_email** (line)

The e-mail address of the person specified in `_software.contact_author`.

Example: 'bourne@sdsc.edu'. [software]

**\_software.date** (line)

The date the software was released.

Examples: '1991-10-01', '1990-04-30'. [software]

**\_software.dependencies** (line)

Any prerequisite software required to run `_software.name`.

Example: 'PDBlib class library'. [software]

**\_software.description** (line)

Description of the software.

Example: 'Uses method of restrained least squares'. [software]

**\_software.hardware** (line)

The hardware upon which the software was run.

Examples: 'Sun Sparc 10 model 41', 'Dec Alpha 3000 model 500S',  
 'Silicon Graphics Elan', 'Compaq PC 486/66'. [software]

**\_software.language** (uline)

The major computing language in which the software is coded.

The data value must be one of the following:

```
Ada
assembler
Awk
Basic
C++
C/C++
C
csh
Fortran
Fortran_77
'Fortran 77'
'Fortran 90'
Java
ksh
Pascal
Perl
Python
sh
Tcl
Other [software]
```

**\_software.location** (line)

The URL for an Internet address at which details of the software can be found.

Examples:  
 'http://rosebud.sdsc.edu/projects/pb/IUCr/software.html',  
 'ftp://ftp.sdsc.edu/pub/sdsc/biology/'. [software]

**\_software.mods** (line)

Any noteworthy modifications to the base software, if applicable.

Example: 'Added support for space group F432'. [software]

**\* \_software.name** (text)

The name of the software.

Examples: 'Merlot', 'O', 'Xengen', 'X-plor'. [software]

**\_software.os** (text)

The name of the operating system under which the software runs.

Examples: 'Ultrinsic', 'OpenVMS', 'DOS', 'Windows 95', 'Windows NT', 'Irix',  
 'HPUX', 'DEC Unix'. [software]

**\_software.os\_version** (text)

The version of the operating system under which the software runs.

Examples: '3.1', '4.2.1'. [software]

**software.type** (uline)  
 The classification of the software according to the most common types.  
 The data value must be one of the following:

program	individual program with limited functionality
library	used by a program at load time
package	collections of programs with multiple functionality
filter	filters input and output streams
jiffy	short, simple program
other	all other kinds of software

[software]

\* **software.version** (line)  
 The version of the software.  
 Examples: 'v1.0', 'beta', '3.1-2', 'unknown'. [software]

<b>SPACE_GROUP</b>	
Contains all the data items that refer to the space group as a whole, such as its name or crystal system. They may be looped, for example, in a list of space groups and their properties. Only a subset of the SPACE_GROUP category items appear in this dictionary. The remainder are found in the symmetry CIF dictionary. Space-group types are identified by their number as given in <i>International Tables for Crystallography</i> Vol. A. Specific settings of the space groups can be identified either by their Hall symbol or by specifying their symmetry operations. The commonly used Hermann–Mauguin symbol determines the space-group type uniquely but several different Hermann–Mauguin symbols may refer to the same space-group type. A Hermann–Mauguin symbol contains information on the choice of the basis, but not on the choice of origin. Different formats for the Hermann–Mauguin symbol are found in the symmetry CIF dictionary.	
Category key(s): <code>_space_group.id</code>	
<i>Example 1 – the monoclinic space group No. 15 with unique axis b.</i>	
<code>_space_group.id</code>	1
<code>_space_group.name_H-M_alt</code>	'C 2/c'
<code>_space_group.IT_number</code>	15
<code>_space_group.name_Hall</code>	'-C 2yc'
<code>_space_group.crystal_system</code>	monoclinic

**space\_group.crystal\_system** (code)  
`space_group_crystal_system` (cif\_core.dic 2.3)  
 The name of the system of geometric crystal classes of space groups (crystal system) to which the space group belongs. Note that rhombohedral space groups belong to the trigonal system.  
 Related item: `symmetry.cell_setting` (alternate).  
 The data value must be one of the following:

triclinic	
monoclinic	
orthorhombic	
tetragonal	
trigonal	
hexagonal	
cubic	[space_group]

\* **space\_group.id** (code)  
`space_group_id` (cif\_core.dic 2.3)  
 This is the unique identifier for the SPACE\_GROUP category.  
 [space\_group]

**space\_group.IT\_number** (int)  
`space_group_IT_number` (cif\_core.dic 2.3)  
 The number as assigned in *International Tables for Crystallography* Vol. A, specifying the proper affine class (i.e. the orientation-preserving affine class) of space groups (crystallographic space-group type) to which the space group belongs. This number defines the space-group type but not the coordinate system in which it is expressed.  
 The permitted range is [1, 230].  
 Related item: `symmetry.Int_Tables_number` (alternate). [space\_group]

**space\_group.name\_H-M\_alt** (line)  
`space_group_name_H-M_alt` (cif\_core.dic 2.3)  
`space_group_name_H-M_alt` allows any Hermann–Mauguin symbol to be given. The way in which this item is used is determined by the user and in general is not intended to be interpreted by computer. It may, for example, be used to give one of the extended Hermann–Mauguin symbols given in Table 4.3.2.1 of *International Tables for Crystallography* Vol. A (2002) or a Hermann–Mauguin symbol for a conventional or unconventional setting. Each component of the space-group name is separated by a space or an underscore. The use of a space is strongly recommended. The underscore is only retained because it was used in old CIFs. It should not be used in new CIFs. Subscripts should appear without special symbols. Bars should be given as negative signs before the numbers to which they apply. The commonly used Hermann–Mauguin symbol determines the space-group type uniquely but a given space-group type may be described by more than one Hermann–Mauguin symbol. The space-group type is best described using `space_group.IT_number`. The Hermann–Mauguin symbol may contain information on the choice of basis, but not on the choice of origin. To define the setting uniquely, use `space_group.name_Hall` or list the symmetry operations.  
 Related item: `symmetry.space_group_name_H-M` (alternate).  
 Example:  

```

; loop_
  _space_group_name_H-M_alt
  'C m c m'
  'C 2/c 2/m 21/m'
  'A m a m'
; (three examples for space group No. 63)
    
```

 [space\_group]

**space\_group.name\_Hall** (line)  
`space_group_name_Hall` (cif\_core.dic 2.3)  
 Space-group symbol defined by Hall. Each component of the space-group name is separated by a space or an underscore. The use of a space is strongly recommended. The underscore is only retained because it was used in old CIFs. It should not be used in new CIFs. `space_group.name_Hall` uniquely defines the space group and its reference to a particular coordinate system.  
 Reference: Hall, S. R. (1981). *Acta Cryst.* **A37**, 517–525; erratum (1981), **A37**, 921. [See also *International Tables for Crystallography* Vol. B (2001), Chapter 1.4, Appendix 1.4.2.]  
 Related item: `symmetry.space_group_name_Hall` (alternate).  
 Examples: 'P 2c -2ac' (equivalent to *Pca2<sub>1</sub>*), '-I 4bd 2ab 3' (equivalent to *Ia3d*).  
 [space\_group]

**SPACE\_GROUP\_SYMOP**

Contains information about the symmetry operations of the space group.

Category key(s): `_space_group_symop.id`

Example 1 – The symmetry operations for the space group  $P2_1/c$ .

```
loop_
  _space_group_symop.id
  _space_group_symop.operation_xyz
  1   x, y, z
  2  -x, -y, -z
  3  -x, 1/2+y, 1/2-z
  4   x, 1/2-y, 1/2+z
```

\* `_space_group_symop.id` (code)

`_space_group_symop.id` (cif\_core.dic 2.3)

An arbitrary identifier that uniquely labels each symmetry operation in the list.

Related item: `_symmetry_equiv.id` (alternate). [space\_group\_symop]

`_space_group_symop.operation_xyz` (line)

`_space_group_symop.operation_xyz` (cif\_core.dic 2.3)

A parsable string giving one of the symmetry operations of the space group in algebraic form. If  $W$  is a matrix representation of the rotational part of the symmetry operation defined by the positions and signs of  $x$ ,  $y$  and  $z$ , and  $w$  is a column of translations defined by fractions, an equivalent position  $x'$  is generated from a given position  $x$  by

$$x' = Wx + w.$$

When a list of symmetry operations is given, it must contain a complete set of coordinate representatives which generates all the operations of the space group by the addition of all primitive translations of the space group. Such representatives are to be found as the coordinates of the general-equivalent position in *International Tables for Crystallography* Vol. A (2002), to which it is necessary to add any centring translations shown above the general-equivalent position. That is to say, it is necessary to list explicitly all the symmetry operations required to generate all the atoms in the unit cell defined by the setting used.

Related item: `_symmetry_equiv.pos_as_xyz` (alternate).

Example: 'x, 1/2-y, 1/2+z' (glide reflection through the plane (x, 1/4, z), with glide vector (1/2)c). [space\_group\_symop]

`_space_group_symop.sg_id` (code)

`_space_group_symop.sg_id` (cif\_core.dic 2.3)

This must match a particular value of `_space_group.id`, allowing the symmetry operation to be identified with a particular space group.

[space\_group\_symop]

**STRUCT**

Data items in the STRUCT category record details about the description of the crystallographic structure.

Category group(s): `inclusive_group`  
`struct_group`

Category key(s): `_struct.entry_id`

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
_struct.entry_id      '5HVP'
_struct.title
; HIV-1 protease complex with acetyl-pepstatin
;
```

\* `_struct.entry_id`

This data item is a pointer to `_entry.id` in the ENTRY category.

`_struct.title` (text)

A title for the data block. The author should attempt to convey the essence of the structure archived in the CIF in the title, and to distinguish this structural result from others.

Examples: '5'-D(\* (I) CP\*CP\*GP\*G) -3', 'T4 lysozyme mutant - S32A',

'hen egg white lysozyme at -30 degrees C',

'quail egg white lysozyme at 2 atmospheres'.

[struct]

**STRUCT\_ASYM**

Data items in the STRUCT\_ASYM category record details about the structural elements in the asymmetric unit.

Category group(s): `inclusive_group`  
`struct_group`

Category key(s): `_struct_asym.id`

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
loop_
  _struct_asym.id
  _struct_asym.entity_id
  _struct_asym.details
  A 1 'one monomer of the dimeric enzyme'
  B 1 'one monomer of the dimeric enzyme'
  C 2 'one partially occupied position for the inhibitor'
  D 2 'one partially occupied position for the inhibitor'
```

`_struct_asym.details` (text)

A description of special aspects of this portion of the contents of the asymmetric unit.

Example:

```
; The drug binds to this enzyme in two roughly twofold
symmetric modes. Hence this biological unit (3) is roughly
twofold symmetric to biological unit (2). Disorder in the
protein chain indicated with alternative ID 2 should be used
with this biological unit.
```

[struct\_asym]

\* `_struct_asym.entity_id`

This data item is a pointer to `_entity.id` in the ENTITY category.

\* `_struct_asym.id` (code)

The value of `_struct_asym.id` must uniquely identify a record in the STRUCT\_ASYM list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

```
_atom_site.label_asym_id,
_struct_biol_gen.asym_id,
_geom_angle.atom_site_label_asym_id 1,
_geom_angle.atom_site_label_asym_id 2,
_geom_angle.atom_site_label_asym_id 3,
_geom_bond.atom_site_label_asym_id 1,
_geom_bond.atom_site_label_asym_id 2,
_geom_contact.atom_site_label_asym_id 1,
_geom_contact.atom_site_label_asym_id 2,
_geom_hbond.atom_site_label_asym_id A,
_geom_hbond.atom_site_label_asym_id D,
_geom_hbond.atom_site_label_asym_id H,
_geom_torsion.atom_site_label_asym_id 1,
_geom_torsion.atom_site_label_asym_id 2,
_geom_torsion.atom_site_label_asym_id 3,
_geom_torsion.atom_site_label_asym_id 4,
_struct_conf.beg_label_asym_id,
_struct_conf.end_label_asym_id,
_struct_conn.ptnr1_label_asym_id,
```

```

_struct_conn.ptnr2_label_asym_id,
_struct_mon_nucl.label_asym_id,
_struct_mon_prot.label_asym_id,
_struct_mon_prot_cis.label_asym_id,
_struct_ncs_dom_lim.beg_label_asym_id,
_struct_ncs_dom_lim.end_label_asym_id,
_struct_sheet_range.beg_label_asym_id,
_struct_sheet_range.end_label_asym_id,
_struct_site_gen.label_asym_id.

```

Examples: '1', 'A', '2B3'.

[struct\_asym]

### STRUCT\_BIOL

Data items in the STRUCT\_BIOL category record details about the structural elements that form each structure of biological significance. A given crystal structure may contain many different biological structures. A given structural component in the asymmetric unit may be part of more than one biological unit. A given biological structure may involve crystallographic symmetry. For instance, in a structure of a lysozyme-FAB structure, the light and heavy-chain components of the FAB could be one biological unit, while the two chains of the FAB and the lysozyme could constitute a second biological unit.

Category group(s): inclusive\_group  
struct\_group

Category key(s): \_struct\_biol.id

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```

loop_
_struct_biol.id
_struct_biol.details
  1
; significant deviations from twofold symmetry exist in this
  dimeric enzyme
;
  2
; The drug binds to this enzyme in two roughly twofold
  symmetric modes. Hence this biological unit (2) is roughly
  twofold symmetric to biological unit (3). Disorder in the
  protein chain indicated with alternative ID 1 should be
  used with this biological unit.
;
  3
; The drug binds to this enzyme in two roughly twofold
  symmetric modes. Hence this biological unit (3) is roughly
  twofold symmetric to biological unit (2). Disorder in the
  protein chain indicated with alternative ID 2 should be
  used with this biological unit.
;

```

### \_struct\_biol.details

(text)

A description of special aspects of the biological unit.

Example:

```

; The drug binds to this enzyme in two roughly twofold
symmetric modes. Hence this biological unit (3) is roughly
twofold symmetric to biological unit (2). Disorder in the
protein chain indicated with alternative ID 2 should be used
with this biological unit.

```

;

[struct\_biol]

### \*\_struct\_biol.id

(line)

The value of \_struct\_biol.id must uniquely identify a record in the STRUCT\_BIOL list. Note that this item need not be a number; it can be any unique identifier.

*The following item(s) have an equivalent role in their respective categories:*

```

_struct_biol_gen.biol_id,
_struct_biol_keywords.biol_id,
_struct_biol_view.biol_id,
_struct_ref.biol_id.

```

[struct\_biol]

### STRUCT\_BIOL\_GEN

Data items in the STRUCT\_BIOL\_GEN category record details about the generation of each biological unit. The STRUCT\_BIOL\_GEN data items provide the specifications of the components that constitute that biological unit, which may include symmetry elements.

Category group(s): inclusive\_group  
struct\_group

Category key(s): \_struct\_biol\_gen.biol\_id  
\_struct\_biol\_gen.asym\_id  
\_struct\_biol\_gen.symmetry

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```

loop_
_struct_biol_gen.biol_id
_struct_biol_gen.asym_id
_struct_biol_gen.symmetry
  1 A 1_555
  1 B 1_555
  2 A 1_555
  2 B 1_555
  2 C 1_555
  3 A 1_555
  3 B 1_555
  3 D 1_555

```

### \*\_struct\_biol\_gen.asym\_id

This data item is a pointer to \_struct\_asym.id in the STRUCT\_ASYM category.

### \*\_struct\_biol\_gen.biol\_id

This data item is a pointer to \_struct\_biol.id in the STRUCT\_BIOL category.

### \_struct\_biol\_gen.details

(text)

A description of special aspects of the symmetry generation of this portion of the biological structure.

Example:

```

; The zinc atom lies on a special position; application of
symmetry elements to generate the insulin hexamer will
generate excess zinc atoms, which must be removed by hand.

```

;

[struct\_biol\_gen]

### \*\_struct\_biol\_gen.symmetry

(symop)

Describes the symmetry operation that should be applied to the atom set specified by \_struct\_biol\_gen.asym\_id to generate a portion of the biological structure.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied), '7-645' (7th symmetry position: +a on x, -b on y).

[struct\_biol\_gen]

### STRUCT\_BIOL\_KEYWORDS

Data items in the STRUCT\_BIOL\_KEYWORDS category record keywords that describe each biological unit.

Category group(s): inclusive\_group  
struct\_group

Category key(s): \_struct\_biol\_keywords.biol\_id  
\_struct\_biol\_keywords.text

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
loop_
_struct_biol_keywords.biol_id
_struct_biol_keywords.text
  1 'aspartyl-protease'
  1 'aspartic-protease'
  1 'acid-protease'
  1 'aspartyl-proteinase'
  1 'aspartic-proteinase'
  1 'acid-proteinase'
  1 'enzyme'
  1 'protease'
  1 'proteinase'
  1 'dimer'
  2 'drug-enzyme complex'
  2 'inhibitor-enzyme complex'
  2 'drug-protease complex'
  2 'inhibitor-protease complex'
  3 'drug-enzyme complex'
  3 'inhibitor-enzyme complex'
  3 'drug-protease complex'
  3 'inhibitor-protease complex'
```

\* **\_struct\_biol\_keywords.biol\_id**

This data item is a pointer to `_struct_biol.id` in the STRUCT\_BIOL category.

\* **\_struct\_biol\_keywords.text**

Keywords describing this biological entity.

Examples: 'antibody', 'antigen', 'enzyme', 'cytokine', 'tRNA'.

[struct\_biol\_keywords]

### STRUCT\_BIOL\_VIEW

Data items in the STRUCT\_BIOL\_VIEW category record details about how to draw and annotate an informative view of the biological structure.

Category group(s): `inclusive_group`  
`struct_group`

Category key(s): `_struct_biol_view.biol_id`  
`_struct_biol_view.id`

Example 1 – based on NDB structure GDL001 by Coll, Aymami, Van Der Marel, Van Boom, Rich & Wang [Biochemistry, (1989), 28, 310–320].

```
_struct_biol_view.biol_id      c1
_struct_biol_view.id          1
_struct_biol_view.rot_matrix[1][1]  0.132
_struct_biol_view.rot_matrix[1][2]  0.922
_struct_biol_view.rot_matrix[1][3] -0.363
_struct_biol_view.rot_matrix[2][1]  0.131
_struct_biol_view.rot_matrix[2][2] -0.380
_struct_biol_view.rot_matrix[2][3] -0.916
_struct_biol_view.rot_matrix[3][1] -0.982
_struct_biol_view.rot_matrix[3][2]  0.073
_struct_biol_view.rot_matrix[3][3] -0.172
_struct_biol_view.details
; This view highlights the ATAT-Netropsin interaction in the
  DNA-drug complex.
;
```

\* **\_struct\_biol\_view.biol\_id**

This data item is a pointer to `_struct_biol.id` in the STRUCT\_BIOL category.

**\_struct\_biol\_view.details**

(text)

A description of special aspects of this view of the biological structure. This data item can be used as a figure legend.

Example:

; The enzyme has been oriented with the molecular twofold axis aligned with the horizontal axis of the figure.

[struct\_biol\_view]

\* **\_struct\_biol\_view.id** (line)

The value of `_struct_biol_view.id` must uniquely identify a record in the STRUCT\_BIOL\_VIEW list. Note that this item need not be a number; it can be any unique identifier.

Examples: 'Figure 1', 'unliganded enzyme',

'view down enzyme active site'.

[struct\_biol\_view]

**\_struct\_biol\_view.rot\_matrix[1][1]** (float)

The [1][1] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_BIOL\_GEN category to give a view useful for describing the structure. The conventions used in the rotation are described in `_struct_biol_view.details`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_biol\_view]

**\_struct\_biol\_view.rot\_matrix[1][2]** (float)

The [1][2] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_BIOL\_GEN category to give a view useful for describing the structure. The conventions used in the rotation are described in `_struct_biol_view.details`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_biol\_view]

**\_struct\_biol\_view.rot\_matrix[1][3]** (float)

The [1][3] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_BIOL\_GEN category to give a view useful for describing the structure. The conventions used in the rotation are described in `_struct_biol_view.details`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_biol\_view]

**\_struct\_biol\_view.rot\_matrix[2][1]** (float)

The [2][1] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_BIOL\_GEN category to give a view useful for describing the structure. The conventions used in the rotation are described in `_struct_biol_view.details`.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_biol\_view]

**struct\_biol\_view.rot\_matrix[2][2]** (float)  
 The [2][2] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_BIOL\_GEN category to give a view useful for describing the structure. The conventions used in the rotation are described in struct\_biol\_view.details.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_biol\_view]

**struct\_biol\_view.rot\_matrix[2][3]** (float)  
 The [2][3] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_BIOL\_GEN category to give a view useful for describing the structure. The conventions used in the rotation are described in struct\_biol\_view.details.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_biol\_view]

**struct\_biol\_view.rot\_matrix[3][1]** (float)  
 The [3][1] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_BIOL\_GEN category to give a view useful for describing the structure. The conventions used in the rotation are described in struct\_biol\_view.details.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_biol\_view]

**struct\_biol\_view.rot\_matrix[3][2]** (float)  
 The [3][2] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_BIOL\_GEN category to give a view useful for describing the structure. The conventions used in the rotation are described in struct\_biol\_view.details.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_biol\_view]

**struct\_biol\_view.rot\_matrix[3][3]** (float)  
 The [3][3] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_BIOL\_GEN category to give a view useful for describing the structure. The conventions used in the rotation are described in struct\_biol\_view.details.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_biol\_view]

## STRUCT\_CONF

Data items in the STRUCT\_CONF category record details about the backbone conformation of a segment of polymer. Data items in the STRUCT\_CONF\_TYPE category define the criteria used to identify the backbone conformations.

Category group(s): inclusive\_group  
struct\_group  
 Category key(s): struct\_conf.id

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
_struct_conf.id
_struct_conf.conf_type_id
_struct_conf.beg_label_comp_id
_struct_conf.beg_label_asym_id
_struct_conf.beg_label_seq_id
_struct_conf.end_label_comp_id
_struct_conf.end_label_asym_id
_struct_conf.end_label_seq_id
_struct_conf.details
HELX1  HELX_RH_AL_P  ARG  A   87  GLN  A   92  .
HELX2  HELX_RH_AL_P  ARG  B  287  GLN  B  292  .
STRN1  STRN          PRO  A    1  LEU  A    5  .
STRN2  STRN          CYS  B  295  PHE  B  299  .
STRN3  STRN          CYS  A   95  PHE  A  299  .
STRN4  STRN          PRO  B  201  LEU  B  205  .
# - - - data truncated for brevity - - -
TURN1  TURN_TY1P_P    ILE  A   15  GLN  A   18  .
TURN2  TURN_TY2_P     GLY  A   49  GLY  A   52  .
TURN3  TURN_TY1P_P    ILE  A   55  HIS  A   69  .
TURN4  TURN_TY1_P     THR  A   91  GLY  A   94  .
# - - - data truncated for brevity - - -
```

**struct\_conf.beg\_auth\_asym\_id**

A component of the identifier for the residue at which the conformation segment begins. This data item is a pointer to atom\_site.auth\_asym\_id in the ATOM\_SITE category.

**struct\_conf.beg\_auth\_comp\_id**

A component of the identifier for the residue at which the conformation segment begins. This data item is a pointer to atom\_site.auth\_comp\_id in the ATOM\_SITE category.

**struct\_conf.beg\_auth\_seq\_id**

A component of the identifier for the residue at which the conformation segment begins. This data item is a pointer to atom\_site.auth\_seq\_id in the ATOM\_SITE category.

**\* struct\_conf.beg\_label\_asym\_id**

A component of the identifier for the residue at which the conformation segment begins. This data item is a pointer to atom\_site.label\_asym\_id in the ATOM\_SITE category.

**\* struct\_conf.beg\_label\_comp\_id**

A component of the identifier for the residue at which the conformation segment begins. This data item is a pointer to atom\_site.label\_comp\_id in the ATOM\_SITE category.

**\* struct\_conf.beg\_label\_seq\_id**

A component of the identifier for the residue at which the conformation segment begins. This data item is a pointer to atom\_site.label\_seq\_id in the ATOM\_SITE category.

**\* struct\_conf.conf\_type\_id**

This data item is a pointer to struct\_conf\_type.id in the STRUCT\_CONF\_TYPE category.

**struct\_conf.details**

(text)  
 A description of special aspects of the conformation assignment.

[struct\_conf]

**\_\_struct\_conf.end\_auth\_asym\_id**

A component of the identifier for the residue at which the conformation segment ends. This data item is a pointer to `__atom_site.auth_asym_id` in the ATOM\_SITE category.

**\_\_struct\_conf.end\_auth\_comp\_id**

A component of the identifier for the residue at which the conformation segment ends. This data item is a pointer to `__atom_site.auth_comp_id` in the ATOM\_SITE category.

**\_\_struct\_conf.end\_auth\_seq\_id**

A component of the identifier for the residue at which the conformation segment ends. This data item is a pointer to `__atom_site.auth_seq_id` in the ATOM\_SITE category.

**\*\_\_struct\_conf.end\_label\_asym\_id**

A component of the identifier for the residue at which the conformation segment ends. This data item is a pointer to `__atom_site.label_asym_id` in the ATOM\_SITE category.

**\*\_\_struct\_conf.end\_label\_comp\_id**

A component of the identifier for the residue at which the conformation segment ends. This data item is a pointer to `__atom_site.label_comp_id` in the ATOM\_SITE category.

**\*\_\_struct\_conf.end\_label\_seq\_id**

A component of the identifier for the residue at which the conformation segment ends. This data item is a pointer to `__atom_site.label_seq_id` in the ATOM\_SITE category.

**\*\_\_struct\_conf.id***(code)*

The value of `__struct_conf.id` must uniquely identify a record in the STRUCT\_CONF list. Note that this item need not be a number; it can be any unique identifier.

[struct\_conf]

**STRUCT\_CONF\_TYPE**

Data items in the STRUCT\_CONF\_TYPE category record details about the criteria used to identify backbone conformations of a segment of polymer.

Category group(s): `inclusive_group`

`struct_group`

Category key(s): `__struct_conf_type.id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

`loop_`

`__struct_conf_type.id`

`__struct_conf_type.criteria`

`__struct_conf_type.reference`

```
HELX_RH_AL_P  'author judgement' .
STRN          'author judgement' .
TURN_TY1_P   'author judgement' .
TURN_TY1P_P  'author judgement' .
TURN_TY2_P   'author judgement' .
TURN_TY2P_P  'author judgement' .
```

**\*\_\_struct\_conf\_type.id***(ucode)*

The descriptor that categorizes the type of the conformation of the backbone of the polymer (whether protein or nucleic acid). Explicit values for the torsion angles that define each conformation are not given here, but it is expected that the author would provide such information in either the `__struct_conf_type.criteria` or `__struct_conf_type.reference` data items, or both.

The following item(s) have an equivalent role in their respective categories:

**\_\_struct\_conf.conf\_type\_id**

The data value must be one of the following:

HELX_P	helix with handedness and type not specified (protein)
HELX_OT_P	helix with handedness and type that do not conform to an accepted category (protein)
HELX_RH_P	right-handed helix with type not specified (protein)
HELX_RH_OT_P	right-handed helix with type that does not conform to an accepted category (protein)
HELX_RH_AL_P	right-handed $\alpha$ helix (protein)
HELX_RH_GA_P	right-handed $\gamma$ helix (protein)
HELX_RH_OM_P	right-handed $\omega$ helix (protein)
HELX_RH_PI_P	right-handed $\pi$ helix (protein)
HELX_RH_27_P	right-handed 2–7 helix (protein)
HELX_RH_3T_P	right-handed 3–10 helix (protein)
HELX_RH_PP_P	right-handed polyproline helix (protein)
HELX_LH_P	left-handed helix with type not specified (protein)
HELX_LH_OT_P	left-handed helix with type that does not conform to an accepted category (protein)
HELX_LH_AL_P	left-handed $\alpha$ helix (protein)
HELX_LH_GA_P	left-handed $\gamma$ helix (protein)
HELX_LH_OM_P	left-handed $\omega$ helix (protein)
HELX_LH_PI_P	left-handed $\pi$ helix (protein)
HELX_LH_27_P	left-handed 2–7 helix (protein)
HELX_LH_3T_P	left-handed 3–10 helix (protein)
HELX_LH_PP_P	left-handed polyproline helix (protein)
HELX_N	helix with handedness and type not specified (nucleic acid)
HELX_OT_N	helix with handedness and type that do not conform to an accepted category (nucleic acid)
HELX_RH_N	right-handed helix with type not specified (nucleic acid)
HELX_RH_OT_N	right-handed helix with type that does not conform to an accepted category (nucleic acid)
HELX_RH_A_N	right-handed A helix (nucleic acid)
HELX_RH_B_N	right-handed B helix (nucleic acid)
HELX_RH_Z_N	right-handed Z helix (nucleic acid)
HELX_LH_N	left-handed helix with type not specified (nucleic acid)
HELX_LH_OT_N	left-handed helix with type that does not conform to an accepted category (nucleic acid)
HELX_LH_A_N	left-handed A helix (nucleic acid)
HELX_LH_B_N	left-handed B helix (nucleic acid)
HELX_LH_Z_N	left-handed Z helix (nucleic acid)
TURN_P	turn with type not specified (protein)
TURN_OT_P	turn with type that does not conform to an accepted category (protein)
TURN_TY1_P	type I turn (protein)
TURN_TY1P_P	type I' turn (protein)
TURN_TY2_P	type II turn (protein)
TURN_TY2P_P	type II' turn (protein)
TURN_TY3_P	type III turn (protein)
TURN_TY3P_P	type III' turn (protein)
STRN	$\beta$ strand (protein)

[struct\_conf\_type]

**\_\_struct\_conf\_type.criteria***(text)*

The criteria used to assign this conformation type.

Examples: 'author judgement', 'phi=54-74, psi=30-50'.

[struct\_conf\_type]

**\_\_struct\_conf\_type.reference***(text)*

A literature reference that defines the criteria used to assign this conformation type and subtype.

[struct\_conf\_type]

## STRUCT\_CONN

Data items in the STRUCT\_CONN category record details about the connections between portions of the structure. These can be hydrogen bonds, salt bridges, disulfide bridges and so on. The STRUCT\_CONN\_TYPE records define the criteria used to identify these connections.

Category group(s): `inclusive_group`  
`struct_group`

Category key(s): `_struct_conn.id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
_struct_conn.id
_struct_conn.conn_type_id
_struct_conn.ptnr1_label_comp_id
_struct_conn.ptnr1_label_asym_id
_struct_conn.ptnr1_label_seq_id
_struct_conn.ptnr1_label_atom_id
_struct_conn.ptnr1_role
_struct_conn.ptnr1_symmetry
_struct_conn.ptnr2_label_comp_id
_struct_conn.ptnr2_label_asym_id
_struct_conn.ptnr2_label_seq_id
_struct_conn.ptnr2_label_atom_id
_struct_conn.ptnr2_role
_struct_conn.ptnr2_symmetry
_struct_conn.details
C1 saltbr ARG A 87 NZ1 positive 1_555 GLU A 92 OE1
   negative 1_555 .
C2 hydrog ARG B 287 N donor 1_555 GLY B 292 O
   acceptor 1_555 .
# - - - - data truncated for brevity - - - -
```

## \*\_struct\_conn.conn\_type\_id

This data item is a pointer to `_struct_conn_type.id` in the STRUCT\_CONN\_TYPE category.

`_struct_conn.details`

(text)

A description of special aspects of the connection.

Example: 'disulfide bridge C-S-S-C is highly distorted'.

[struct\_conn]

## \*\_struct\_conn.id

(code)

The value of `_struct_conn.id` must uniquely identify a record in the STRUCT\_CONN list. Note that this item need not be a number; it can be any unique identifier.

[struct\_conn]

`_struct_conn.ptnr1_auth_asym_id`

A component of the identifier for partner 1 of the structure connection. This data item is a pointer to `_atom_site.auth_asym_id` in the ATOM\_SITE category.

`_struct_conn.ptnr1_auth_atom_id`

A component of the identifier for partner 1 of the structure connection. This data item is a pointer to `_atom_site.auth_atom_id` in the ATOM\_SITE category.

`_struct_conn.ptnr1_auth_comp_id`

A component of the identifier for partner 1 of the structure connection. This data item is a pointer to `_atom_site.auth_comp_id` in the ATOM\_SITE category.

`_struct_conn.ptnr1_auth_seq_id`

A component of the identifier for partner 1 of the structure connection. This data item is a pointer to `_atom_site.auth_seq_id` in the ATOM\_SITE category.

`_struct_conn.ptnr1_label_alt_id`

A component of the identifier for partner 1 of the structure connection. This data item is a pointer to `_atom_sites_alt.id` in the ATOM\_SITES\_ALT category.

## \*\_struct\_conn.ptnr1\_label\_asym\_id

A component of the identifier for partner 1 of the structure connection. This data item is a pointer to `_atom_site.label_asym_id` in the ATOM\_SITE category.

## \*\_struct\_conn.ptnr1\_label\_atom\_id

A component of the identifier for partner 1 of the structure connection. This data item is a pointer to `_chem_comp_atom.atom_id` in the CHEM\_COMP\_ATOM category.

## \*\_struct\_conn.ptnr1\_label\_comp\_id

A component of the identifier for partner 1 of the structure connection. This data item is a pointer to `_atom_site.label_comp_id` in the ATOM\_SITE category.

## \*\_struct\_conn.ptnr1\_label\_seq\_id

A component of the identifier for partner 1 of the structure connection. This data item is a pointer to `_atom_site.label_seq_id` in the ATOM\_SITE category.

`_struct_conn.ptnr1_role`

(uline)

The chemical or structural role of the first partner in the structure connection.

Examples: 'donor', 'acceptor', 'negative', 'positive', 'metal',  
'metal coordination'.

[struct\_conn]

`_struct_conn.ptnr1_symmetry`

(symop)

Describes the symmetry operation that should be applied to the atom set specified by `_struct_conn.ptnr1_label*` to generate the first partner in the structure connection.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),  
'7\_645' (7th symmetry position: +a on x, -b on y).

[struct\_conn]

`_struct_conn.ptnr2_auth_asym_id`

A component of the identifier for partner 2 of the structure connection. This data item is a pointer to `_atom_site.auth_asym_id` in the ATOM\_SITE category.

`_struct_conn.ptnr2_auth_atom_id`

A component of the identifier for partner 2 of the structure connection. This data item is a pointer to `_atom_site.auth_atom_id` in the ATOM\_SITE category.

`_struct_conn.ptnr2_auth_comp_id`

A component of the identifier for partner 2 of the structure connection. This data item is a pointer to `_atom_site.auth_comp_id` in the ATOM\_SITE category.

`_struct_conn.ptnr2_auth_seq_id`

A component of the identifier for partner 2 of the structure connection. This data item is a pointer to `_atom_site.auth_seq_id` in the ATOM\_SITE category.

**\_struct\_conn.ptnr2\_label\_alt\_id**

A component of the identifier for partner 2 of the structure connection. This data item is a pointer to `_atom_sites_alt.id` in the `ATOM_SITES_ALT` category.

**\*\_struct\_conn.ptnr2\_label\_asym\_id**

A component of the identifier for partner 2 of the structure connection. This data item is a pointer to `_atom_site.label_asym_id` in the `ATOM_SITE` category.

**\*\_struct\_conn.ptnr2\_label\_atom\_id**

A component of the identifier for partner 2 of the structure connection. This data item is a pointer to `_chem_comp_atom.atom_id` in the `CHEM_COMP_ATOM` category.

**\*\_struct\_conn.ptnr2\_label\_comp\_id**

A component of the identifier for partner 2 of the structure connection. This data item is a pointer to `_atom_site.label_comp_id` in the `ATOM_SITE` category.

**\*\_struct\_conn.ptnr2\_label\_seq\_id**

A component of the identifier for partner 2 of the structure connection. This data item is a pointer to `_atom_site.label_seq_id` in the `ATOM_SITE` category.

**\_struct\_conn.ptnr2\_role** *(uline)*

The chemical or structural role of the second partner in the structure connection.

Examples: 'donor', 'acceptor', 'negative', 'positive', 'metal', 'metal coordination'. [struct\_conn]

**\_struct\_conn.ptnr2\_symmetry** *(symop)*

Describes the symmetry operation that should be applied to the atom set specified by `_struct_conn.ptnr2_label*` to generate the second partner in the structure connection.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied), '7\_645' (7th symmetry position: +a on x, -b on y). [struct\_conn]

**STRUCT\_CONN\_TYPE**

Data items in the `STRUCT_CONN_TYPE` category record details about the criteria used to identify interactions between portions of the structure.

Category group(s): `inclusive_group`  
`struct_group`  
Category key(s): `_struct_conn_type.id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _struct_conn_type.id
  _struct_conn_type.criteria
  _struct_conn_type.reference
  saltbr
  'negative to positive distance > 2.5 \ %A, < 3.2 \ %A' .
  hydrog
  'NO distance > 2.5\ %A, < 3.5\ %A, NOC angle < 120 degrees' .
```

**\_struct\_conn\_type.criteria** *(text)*

The criteria used to define the interaction.

Examples: 'O to N distance > 2.5 \ %A, < 3.2 \ %A', 'authors judgement'. [struct\_conn\_type]

**\*\_struct\_conn\_type.id** *(ucode)*

The chemical or structural type of the interaction.

The following item(s) have an equivalent role in their respective categories:

`_struct_conn.conn_type_id`.

The data value must be one of the following:

covale	covalent bond
disulf	disulfide bridge
hydrog	hydrogen bond
metalc	metal coordination
mismat	mismatched base pairs
saltbr	ionic interaction
modres	covalent residue modification
covale_base	covalent modification of a nucleotide base
covale_sugar	covalent modification of a nucleotide sugar
covale_phosphate	covalent modification of a nucleotide phosphate

[struct\_conn\_type]

**\_struct\_conn\_type.reference** *(text)*

A reference that specifies the criteria used to define the interaction.

[struct\_conn\_type]

**STRUCT\_KEYWORDS**

Data items in the `STRUCT_KEYWORDS` category specify keywords that describe the chemical structure in this entry.

Category group(s): `inclusive_group`  
`struct_group`

Category key(s): `_struct_keywords.entry_id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _struct_keywords.entry_id
  _struct_keywords.text
  '5HVP' 'enzyme-inhibitor complex'
  '5HVP' 'aspartyl protease'
  '5HVP' 'structure-based drug design'
  '5HVP' 'static disorder'
```

**\*\_struct\_keywords.entry\_id**

This data item is a pointer to `_entry.id` in the `ENTRY` category.

**\_struct\_keywords.text** *(text)*

Keywords describing this structure.

Examples: 'serine protease', 'inhibited complex', 'high-resolution refinement'. [struct\_keywords]

**STRUCT\_MON\_DETAILS**

Data items in the `STRUCT_MON_DETAILS` category record details about specifics of calculations summarized in data items in the `STRUCT_MON_PROT` and `STRUCT_MON_NUCL` categories. These can include the coefficients used in map calculations, the radii used for including points in a calculation and so on.

Category group(s): `inclusive_group`  
`struct_group`  
Category key(s): `_struct_mon_details.entry_id`

**\*\_struct\_mon\_details.entry\_id**

This data item is a pointer to `_entry.id` in the `ENTRY` category.

**\_struct\_mon\_details.prot\_cis** *(float)*

An ideal *cis* peptide bond would have an  $\omega$  torsion angle of zero. This data item gives the value in degrees by which the observed torsion angle can differ from 0.0 and still be considered *cis*.

Example: '30.0'. [struct\_mon\_details]

**struct\_mon\_details.RSCC** (text)

This data item describes the specifics of the calculations that generated the values given in struct\_mon\_prot.RSCC\_all, struct\_mon\_prot.RSCC\_main and struct\_mon\_prot.RSCC\_side. The coefficients used to calculate the  $p(o)$  and  $p(c)$  maps should be given as well as the criterion for the inclusion of map grid points in the calculation.

Examples:

```
; The map p(o) was calculated with coefficients
2F(o) - F(c) and with phase alpha(c). F(o)
are the observed structure-factor amplitudes,
F(c) are the amplitudes calculated from the
current model and alpha(c) are the phases
calculated from the current model.
The map p(c) was calculated in program O using
a Gaussian distribution function around the
atoms in the current model.
Map grid points within 1.5 A of the
designated atoms were included in the
calculation.
```

```
; The map p(o) was calculated with coefficients
F(o) and with phase alpha(c). F(o) are the
observed structure-factor amplitudes, and
alpha(c) are the phases calculated from the
current model.
The map p(c) was calculated with coefficients
F(c) and with phases alpha(c). F(c) and
alpha(c) are the structure-factor amplitudes
and phases, respectively, calculated from the
current model.
Map grid points within a van der Waals radius
of the designated atoms were included in the
calculation.
```

[struct\_mon\_details]

**struct\_mon\_details.RSR** (text)

This data item describes the specifics of the calculations that generated the values given in struct\_mon\_prot.RSR\_all, struct\_mon\_prot.RSR\_main and struct\_mon\_prot.RSR\_side. The coefficients used to calculate the  $p(o)$  and  $p(c)$  maps should be given as well as the criterion for the inclusion of map grid points in the calculation.

Examples:

```
; The map p(o) was calculated with coefficients
2F(o) - F(c) and with phase alpha(c). F(o)
are the observed structure-factor amplitudes,
F(c) are the amplitudes calculated from the
current model and alpha(c) are the phases
calculated from the current model.
The map p(c) was calculated in program O using
a Gaussian distribution function around the
atoms in the current model.
Map grid points within 1.5 A of the
designated atoms were included in the
calculation.
```

```
; The map p(o) was calculated with coefficients
F(o) and with phase alpha(c). F(o) are the
observed structure-factor amplitudes, and
alpha(c) are the phases calculated from the
current model.
The map p(c) was calculated with coefficients
F(c) and with phases alpha(c). F(c) and
alpha(c) are the structure-factor amplitudes
and phases, respectively, calculated from the
current model.
Map grid points within a van der Waals radius
of the designated atoms were included in the
calculation.
```

[struct\_mon\_details]

**STRUCT\_MON\_NUCL**

Data items in the STRUCT\_MON\_NUCL category record details about structural properties of a nucleic acid when analyzed at the monomer level. Analogous data items for proteins are given in the STRUCT\_MON\_PROT category. For items where the value of the property depends on the method employed to calculate it, details of the method of calculation are given using data items in the STRUCT\_MON\_DETAILS category.

Category group(s): inclusive\_group

struct\_group

Category key(s): struct\_mon\_nucl.label\_alt\_id

struct\_mon\_nucl.label\_asym\_id

struct\_mon\_nucl.label\_comp\_id

struct\_mon\_nucl.label\_seq\_id

Example 1 – based on NDB structure BDL028.

```
loop_
  struct_mon_nucl.label_comp_id
  struct_mon_nucl.label_seq_id
  struct_mon_nucl.label_asym_id
  struct_mon_nucl.label_alt_id
  struct_mon_nucl.alpha
  struct_mon_nucl.beta
  struct_mon_nucl.gamma
  struct_mon_nucl.delta
  struct_mon_nucl.epsilon
  struct_mon_nucl.zeta
C 1 A . . . 29.9 131.9 222.1 174.2
G 2 A . 334.0 130.6 33.1 125.6 167.6 270.9
T 3 A . 258.2 178.7 101.0 114.6 216.6 259.3
# ----- abbreviated list -----
```

**struct\_mon\_nucl.alpha** (float)

The value in degrees of the backbone torsion angle  $\alpha$  (O3'—P—O5'—C5').

[struct\_mon\_nucl]

**struct\_mon\_nucl.auth\_asym\_id**

A component of the identifier for participants in the site. This data item is a pointer to atom\_site.auth\_asym\_id in the ATOM\_SITE category.

**struct\_mon\_nucl.auth\_comp\_id**

A component of the identifier for participants in the site. This data item is a pointer to atom\_site.auth\_comp\_id in the ATOM\_SITE category.

**struct\_mon\_nucl.auth\_seq\_id**

A component of the identifier for participants in the site. This data item is a pointer to atom\_site.auth\_seq\_id in the ATOM\_SITE category.

**struct\_mon\_nucl.beta** (float)

The value in degrees of the backbone torsion angle  $\beta$  (P—O5'—C5'—C4').

[struct\_mon\_nucl]

**struct\_mon\_nucl.chi1** (float)

The value in degrees of the sugar–base torsion angle  $\chi_1$  (O4'—C1'—N1—C2).

[struct\_mon\_nucl]

**struct\_mon\_nucl.chi2** (float)

The value in degrees of the sugar–base torsion angle  $\chi_2$  (O4'—C1'—N9—C4).

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.delta** (float)  
The value in degrees of the backbone torsion angle  $\delta$  (C5'—C4'—C3'—O3').

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.details** (float)  
A description of special aspects of the residue, its conformation, behaviour in refinement, or any other aspect that requires annotation.

Example:

; Part of the phosphodiester backbone not in density.

; [struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.epsilon** (float)  
The value in degrees of the backbone torsion angle  $\varepsilon$  (C4'—C3'—O3'—P).

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.gamma** (float)  
The value in degrees of the backbone torsion angle  $\gamma$  (O5'—C5'—C4'—C3').

[struct\_mon\_nucl]

\* **\_\_struct\_mon\_nucl.label\_alt\_id**

A component of the identifier for participants in the site. This data item is a pointer to `_atom_sites_alt.id` in the `ATOM_SITES_ALT` category.

\* **\_\_struct\_mon\_nucl.label\_asym\_id**

A component of the identifier for participants in the site. This data item is a pointer to `_atom_site.label_asym_id` in the `ATOM_SITE` category.

\* **\_\_struct\_mon\_nucl.label\_comp\_id**

A component of the identifier for participants in the site. This data item is a pointer to `_atom_site.label_comp_id` in the `ATOM_SITE` category.

\* **\_\_struct\_mon\_nucl.label\_seq\_id**

A component of the identifier for participants in the site. This data item is a pointer to `_atom_site.label_seq_id` in the `ATOM_SITE` category.

**\_\_struct\_mon\_nucl.mean\_B\_all** (float)  
The mean value of the isotropic displacement parameter for all atoms in the monomer.

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.mean\_B\_base** (float)  
The mean value of the isotropic displacement parameter for atoms in the base moiety of the nucleic acid monomer.

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.mean\_B\_phos** (float)  
The mean value of the isotropic displacement parameter for atoms in the phosphate moiety of the nucleic acid monomer.

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.mean\_B\_sugar** (float)  
The mean value of the isotropic displacement parameter for atoms in the sugar moiety of the nucleic acid monomer.

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.nu0** (float)  
The value in degrees of the sugar torsion angle  $\nu_0$  (C4'—O4'—C1'—C2').

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.nu1** (float)  
The value in degrees of the sugar torsion angle  $\nu_1$  (O4'—C1'—C2'—C3').

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.nu2** (float)  
The value in degrees of the sugar torsion angle  $\nu_2$  (C1'—C2'—C3'—C4').

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.nu3** (float)  
The value in degrees of the sugar torsion angle  $\nu_3$  (C2'—C3'—C4'—O4').

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.nu4** (float)  
The value in degrees of the sugar torsion angle  $\nu_4$  (C3'—C4'—O4'—C1').

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.P** (float)  
 $P$  is the phase angle of pseudorotation for five-membered rings. For ribose and deoxyribose sugars in nucleic acids

$$P = \arctan \left( \frac{(\tau_4 + \tau_1) - (\tau_3 + \tau_0)}{2\tau_2(\sin 36 + \sin 72)} \right).$$

If  $\tau_2$  is  $< 0$ , then  $P = P + 180^\circ$  (Altona & Sundaralingam, 1972).

Reference: Altona, C. & Sundaralingam, M. (1972). *J. Am. Chem. Soc.* **94**, 8205–8212.

[struct\_mon\_nucl]

**\_\_struct\_mon\_nucl.RSCC\_all** (float)  
The real-space (linear) correlation coefficient RSCC, as described by Jones *et al.* (1991), evaluated over all atoms in the nucleic acid monomer.

$$\text{RSCC} = \frac{\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle| \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|}{(\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle|^2 \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|^2)^{1/2}},$$

where  $p_{\text{obs}}$  = the density in an 'experimental' map,  $p_{\text{calc}}$  = the density in a 'calculated' map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `__struct_mon_details.RSCC`.  $\langle \rangle$  indicates an average and the sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `__struct_mon_details.RSCC`.

Reference: Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.

[struct\_mon\_nucl]

`_struct_mon_nucl.RSCC_base` (float)

The real-space (linear) correlation coefficient RSCC, as described by Jones *et al.* (1991), evaluated over all atoms in the base moiety of the nucleic acid monomer.

$$\text{RSCC} = \frac{\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle| \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|}{(\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle|^2 \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|^2)^{1/2}},$$

where  $p_{\text{obs}}$  = the density in an ‘experimental’ map,  $p_{\text{calc}}$  = the density in a ‘calculated’ map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `_struct_mon_details.RSCC`.  $\langle \rangle$  indicates an average and the sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSCC`.

Reference: Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.

[struct\_mon\_nucl]

`_struct_mon_nucl.RSCC_phos` (float)

The real-space (linear) correlation coefficient RSCC, as described by Jones *et al.* (1991), evaluated over all atoms in the phosphate moiety of the nucleic acid monomer.

$$\text{RSCC} = \frac{\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle| \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|}{(\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle|^2 \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|^2)^{1/2}},$$

where  $p_{\text{obs}}$  = the density in an ‘experimental’ map,  $p_{\text{calc}}$  = the density in a ‘calculated’ map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `_struct_mon_details.RSCC`.  $\langle \rangle$  indicates an average and the sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSCC`.

Reference: Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.

[struct\_mon\_nucl]

`_struct_mon_nucl.RSCC_sugar` (float)

The real-space (linear) correlation coefficient RSCC, as described by Jones *et al.* (1991), evaluated over all atoms in the sugar moiety of the nucleic acid monomer.

$$\text{RSCC} = \frac{\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle| \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|}{(\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle|^2 \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|^2)^{1/2}},$$

where  $p_{\text{obs}}$  = the density in an ‘experimental’ map,  $p_{\text{calc}}$  = the density in a ‘calculated’ map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `_struct_mon_details.RSCC`.  $\langle \rangle$  indicates an average and the sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSCC`.

Reference: Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.

[struct\_mon\_nucl]

`_struct_mon_nucl.RSR_all` (float)

The real-space residual RSR, as described by Brändén & Jones (1990), evaluated over all atoms in the nucleic acid monomer.

$$\text{RSR} = \frac{\sum |p_{\text{obs}} - p_{\text{calc}}|}{\sum |p_{\text{obs}} + p_{\text{calc}}|},$$

where  $p_{\text{obs}}$  = the density in an ‘experimental’ map,  $p_{\text{calc}}$  = the density in a ‘calculated’ map and the sum is taken over the specified grid points. Details of how these maps were calculated

should be given in `_struct_mon_details.RSR`. The sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSR`.

Reference: Brändén, C.-I. & Jones, T. A. (1990). *Nature (London)*, **343**, 687–689.

[struct\_mon\_nucl]

`_struct_mon_nucl.RSR_base` (float)

The real-space residual RSR, as described by Brändén & Jones (1990), evaluated over all atoms in the base moiety of the nucleic acid monomer.

$$\text{RSR} = \frac{\sum |p_{\text{obs}} - p_{\text{calc}}|}{\sum |p_{\text{obs}} + p_{\text{calc}}|},$$

where  $p_{\text{obs}}$  = the density in an ‘experimental’ map,  $p_{\text{calc}}$  = the density in a ‘calculated’ map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `_struct_mon_details.RSR`. The sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSR`.

Reference: Brändén, C.-I. & Jones, T. A. (1990). *Nature (London)*, **343**, 687–689.

[struct\_mon\_nucl]

`_struct_mon_nucl.RSR_phos` (float)

The real-space residual RSR, as described by Brändén & Jones (1990), evaluated over all atoms in the phosphate moiety of the nucleic acid monomer.

$$\text{RSR} = \frac{\sum |p_{\text{obs}} - p_{\text{calc}}|}{\sum |p_{\text{obs}} + p_{\text{calc}}|},$$

where  $p_{\text{obs}}$  = the density in an ‘experimental’ map,  $p_{\text{calc}}$  = the density in a ‘calculated’ map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `_struct_mon_details.RSR`. The sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSR`.

Reference: Brändén, C.-I. & Jones, T. A. (1990). *Nature (London)*, **343**, 687–689.

[struct\_mon\_nucl]

`_struct_mon_nucl.RSR_sugar` (float)

The real-space residual RSR, as described by Brändén & Jones (1990), evaluated over all atoms in the sugar moiety of the nucleic acid monomer.

$$\text{RSR} = \frac{\sum |p_{\text{obs}} - p_{\text{calc}}|}{\sum |p_{\text{obs}} + p_{\text{calc}}|},$$

where  $p_{\text{obs}}$  = the density in an ‘experimental’ map,  $p_{\text{calc}}$  = the density in a ‘calculated’ map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `_struct_mon_details.RSR`. The sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSR`.

Reference: Brändén, C.-I. & Jones, T. A. (1990). *Nature (London)*, **343**, 687–689.

[struct\_mon\_nucl]

**\_struct\_mon\_nucl.tau0** (float)  
The value in degrees of the sugar torsion angle  $\tau_0$  (C4'—O4'—C1'—C2').

[struct\_mon\_nucl]

**\_struct\_mon\_nucl.tau1** (float)  
The value in degrees of the sugar torsion angle  $\tau_1$  (O4'—C1'—C2'—C3').

[struct\_mon\_nucl]

**\_struct\_mon\_nucl.tau2** (float)  
The value in degrees of the sugar torsion angle  $\tau_2$  (C1'—C2'—C3'—C4').

[struct\_mon\_nucl]

**\_struct\_mon\_nucl.tau3** (float)  
The value in degrees of the sugar torsion angle  $\tau_3$  (C2'—C3'—C4'—O4').

[struct\_mon\_nucl]

**\_struct\_mon\_nucl.tau4** (float)  
The value in degrees of the sugar torsion angle  $\tau_4$  (C3'—C4'—O4'—C1').

[struct\_mon\_nucl]

**\_struct\_mon\_nucl.taum** (float)  
The maximum amplitude of puckering. This is derived from the pseudorotation value  $P$  and the torsion angles in the ribose ring.

$$\begin{aligned}\tau_2 &= \tau_m \cos P, \\ \tau_3 &= \tau_m \cos(P + 144), \\ \tau_4 &= \tau_m \cos(P + 288), \\ \tau_0 &= \tau_m \cos(P + 72), \\ \tau_1 &= \tau_m \cos(P + 216).\end{aligned}$$

[struct\_mon\_nucl]

**\_struct\_mon\_nucl.zeta** (float)  
The value in degrees of the backbone torsion angle  $\zeta$  (C3'—O3'—P—O5').

[struct\_mon\_nucl]

### STRUCT\_MON\_PROT

Data items in the STRUCT\_MON\_PROT category record details about structural properties of a protein when analyzed at the monomer level. Analogous data items for nucleic acids are given in the STRUCT\_MON\_NUCL category. For items where the value of the property depends on the method employed to calculate it, details of the method of calculation are given using data items in the STRUCT\_MON\_DETAILS category.

Category group(s): **inclusive\_group**

struct\_group

Category key(s): **\_struct\_mon\_prot.label\_alt\_id**  
**\_struct\_mon\_prot.label\_asym\_id**  
**\_struct\_mon\_prot.label\_comp\_id**  
**\_struct\_mon\_prot.label\_seq\_id**

*Example 1 – based on laboratory records for protein NSI. This example provides details for residue ARG 35.*

```
_struct_mon_prot.label_comp_id  ARG
_struct_mon_prot.label_seq_id   35
_struct_mon_prot.label_asym_id   A
_struct_mon_prot.label_alt_id    .
_struct_mon_prot.chi1           -67.9
_struct_mon_prot.chi2           -174.7
_struct_mon_prot.chi3           -67.7
_struct_mon_prot.chi4           -86.3
_struct_mon_prot.chi5            4.2
_struct_mon_prot.RSCC_all        0.90
_struct_mon_prot.RSR_all         0.18
_struct_mon_prot.mean_B_all      30.0
_struct_mon_prot.mean_B_main     25.0
_struct_mon_prot.mean_B_side     35.1
_struct_mon_prot.omega           180.1
_struct_mon_prot.phi             -60.3
_struct_mon_prot.psi             -46.0
```

**\_struct\_mon\_prot.auth\_asym\_id**  
A component of the identifier for the monomer. This data item is a pointer to **\_atom\_site.auth\_asym\_id** in the ATOM\_SITE category.

**\_struct\_mon\_prot.auth\_comp\_id**  
A component of the identifier for the monomer. This data item is a pointer to **\_atom\_site.auth\_comp\_id** in the ATOM\_SITE category.

**\_struct\_mon\_prot.auth\_seq\_id**  
A component of the identifier for the monomer. This data item is a pointer to **\_atom\_site.auth\_seq\_id** in the ATOM\_SITE category.

**\_struct\_mon\_prot.chi1** (float)  
The value in degrees of the side-chain torsion angle  $\chi_1$ , for those residues containing such an angle.

[struct\_mon\_prot]

**\_struct\_mon\_prot.chi2** (float)  
The value in degrees of the side-chain torsion angle  $\chi_2$ , for those residues containing such an angle.

[struct\_mon\_prot]

**\_struct\_mon\_prot.chi3** (float)  
The value in degrees of the side-chain torsion angle  $\chi_3$ , for those residues containing such an angle.

[struct\_mon\_prot]

**\_struct\_mon\_prot.chi4** (float)  
The value in degrees of the side-chain torsion angle  $\chi_4$ , for those residues containing such an angle.

[struct\_mon\_prot]

**\_struct\_mon\_prot.chi5** (float)  
The value in degrees of the side-chain torsion angle  $\chi_5$ , for those residues containing such an angle.

[struct\_mon\_prot]

**\_struct\_mon\_prot.details** (float)

A description of special aspects of the residue, its conformation, behaviour in refinement, or any other aspect that requires annotation.

Examples: 'very poor density',  
; The side chain of this density may occupy alternative conformations, but alternative conformations were not fit in this model.

;  
; This residue has a close contact with the bound inhibitor, which may account for the nonstandard conformation of the side chain.

;  
[struct\_mon\_prot]

**\*\_struct\_mon\_prot.label\_alt\_id**

A component of the identifier for the monomer. This data item is a pointer to `_atom_sites_alt.id` in the ATOM\_SITES\_ALT category.

**\*\_struct\_mon\_prot.label\_asym\_id**

A component of the identifier for the monomer. This data item is a pointer to `_atom_site.label_asym_id` in the ATOM\_SITE category.

**\*\_struct\_mon\_prot.label\_comp\_id**

A component of the identifier for the monomer. This data item is a pointer to `_atom_site.label_comp_id` in the ATOM\_SITE category.

**\*\_struct\_mon\_prot.label\_seq\_id**

A component of the identifier for the monomer. This data item is a pointer to `_atom_site.label_seq_id` in the ATOM\_SITE category.

**\_struct\_mon\_prot.mean\_B\_all** (float)

The mean value of the isotropic displacement parameter for all atoms in the monomer.

[struct\_mon\_prot]

**\_struct\_mon\_prot.mean\_B\_main** (float)

The mean value of the isotropic displacement parameter for atoms in the main chain of the monomer.

[struct\_mon\_prot]

**\_struct\_mon\_prot.mean\_B\_side** (float)

The mean value of the isotropic displacement parameter for atoms in the side chain of the monomer.

[struct\_mon\_prot]

**\_struct\_mon\_prot.omega** (float)

The value in degrees of the main-chain torsion angle  $\omega$ .

[struct\_mon\_prot]

**\_struct\_mon\_prot.phi** (float)

The value in degrees of the main-chain torsion angle  $\varphi$ .

[struct\_mon\_prot]

**\_struct\_mon\_prot.psi** (float)

The value in degrees of the main-chain torsion angle  $\psi$ .

[struct\_mon\_prot]

**\_struct\_mon\_prot.RSCC\_all** (float)

The real-space (linear) correlation coefficient RSCC, as described by Jones *et al.* (1991), evaluated over all atoms in the monomer.

$$\text{RSCC} = \frac{\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle| \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|}{(\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle|^2 \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|^2)^{1/2}},$$

where  $p_{\text{obs}}$  = the density in an 'experimental' map,  $p_{\text{calc}}$  = the density in a 'calculated' map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `_struct_mon_details.RSCC`.  $\langle \rangle$  indicates an average and the sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSCC`.

Reference: Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.

[struct\_mon\_prot]

**\_struct\_mon\_prot.RSCC\_main** (float)

The real-space (linear) correlation coefficient RSCC, as described by Jones *et al.* (1991), evaluated over all atoms in the main chain of the monomer.

$$\text{RSCC} = \frac{\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle| \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|}{[\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle|^2 \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|^2]^{1/2}},$$

where  $p_{\text{obs}}$  = the density in an 'experimental' map,  $p_{\text{calc}}$  = the density in a 'calculated' map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `_struct_mon_details.RSCC`.  $\langle \rangle$  indicates an average and the sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSCC`.

Reference: Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.

[struct\_mon\_prot]

**\_struct\_mon\_prot.RSCC\_side** (float)

The real-space (linear) correlation coefficient RSCC, as described by Jones *et al.* (1991), evaluated over all atoms in the side chain of the monomer.

$$\text{RSCC} = \frac{\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle| \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|}{[\sum |p_{\text{obs}} - \langle p_{\text{obs}} \rangle|^2 \sum |p_{\text{calc}} - \langle p_{\text{calc}} \rangle|^2]^{1/2}},$$

where  $p_{\text{obs}}$  = the density in an 'experimental' map,  $p_{\text{calc}}$  = the density in a 'calculated' map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `_struct_mon_details.RSCC`.  $\langle \rangle$  indicates an average and the sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSCC`.

Reference: Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* **A47**, 110–119.

[struct\_mon\_prot]

**\_struct\_mon\_prot.RSR\_all** (float)

The real-space residual RSR, as described by Brändén & Jones (1990), evaluated over all atoms in the monomer.

$$\text{RSR} = \frac{\sum |p_{\text{obs}} - p_{\text{calc}}|}{\sum |p_{\text{obs}} + p_{\text{calc}}|},$$

where  $p_{\text{obs}}$  = the density in an 'experimental' map,  $p_{\text{calc}}$  = the density in a 'calculated' map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `_struct_mon_details.RSR`. The sums are taken

over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSR`.

Reference: Brändén, C.-I. & Jones, T. A. (1990). *Nature (London)*, **343**, 687–689.

[struct\_mon\_prot]

`_struct_mon_prot.RSR_main` (float)

The real-space residual RSR, as described by Brändén & Jones (1990), evaluated over all atoms in the main chain of the monomer.

$$RSR = \frac{\sum |p_{obs} - p_{calc}|}{\sum |p_{obs} + p_{calc}|},$$

where  $p_{obs}$  = the density in an ‘experimental’ map,  $p_{calc}$  = the density in a ‘calculated’ map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `_struct_mon_details.RSR`. The sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSR`.

Reference: Brändén, C.-I. & Jones, T. A. (1990). *Nature (London)*, **343**, 687–689.

[struct\_mon\_prot]

`_struct_mon_prot.RSR_side` (float)

The real-space residual RSR, as described by Brändén & Jones (1990), evaluated over all atoms in the side chain of the monomer.

$$RSR = \frac{\sum |p_{obs} - p_{calc}|}{\sum |p_{obs} + p_{calc}|},$$

where  $p_{obs}$  = the density in an ‘experimental’ map,  $p_{calc}$  = the density in a ‘calculated’ map and the sum is taken over the specified grid points. Details of how these maps were calculated should be given in `_struct_mon_details.RSR`. The sums are taken over all map grid points near the relevant atoms. The radius for including grid points in the calculation should also be given in `_struct_mon_details.RSR`.

Reference: Brändén, C.-I. & Jones, T. A. (1990). *Nature (London)*, **343**, 687–689.

[struct\_mon\_prot]

### STRUCT\_MON\_PROT\_CIS

Data items in the STRUCT\_MON\_PROT\_CIS category identify monomers that have been found to have the peptide bond in the *cis* conformation. The criterion used to select residues to be designated as containing *cis* peptide bonds is given in `_struct_mon_details.prot_cis`.

Category group(s): `inclusive_group`  
`struct_group`  
 Category key(s): `_struct_mon_prot_cis.label_alt_id`  
`_struct_mon_prot_cis.label_asym_id`  
`_struct_mon_prot_cis.label_comp_id`  
`_struct_mon_prot_cis.label_seq_id`

Example 1 – based on PDB structure 1ACY of Ghiara, Stura, Stanfield, Profy & Wilson [Science (1994), **264**, 82–85].

```
loop_
_struct_mon_prot_cis.label_comp_id
_struct_mon_prot_cis.label_seq_id
_struct_mon_prot_cis.label_asym_id
_struct_mon_prot_cis.label_alt_id
PRO 8 L .
PRO 77 L .
PRO 95 L .
PRO 141 L .
# ----- abbreviated -----
```

`_struct_mon_prot_cis.auth_asym_id`

A component of the identifier for the monomer. This data item is a pointer to `_atom_site.auth_asym_id` in the ATOM\_SITE category.

`_struct_mon_prot_cis.auth_comp_id`

A component of the identifier for the monomer. This data item is a pointer to `_atom_site.auth_comp_id` in the ATOM\_SITE category.

`_struct_mon_prot_cis.auth_seq_id`

A component of the identifier for the monomer. This data item is a pointer to `_atom_site.auth_seq_id` in the ATOM\_SITE category.

\* `_struct_mon_prot_cis.label_alt_id`

A component of the identifier for the monomer. This data item is a pointer to `_atom_sites_alt.id` in the ATOM\_SITES\_ALT category.

\* `_struct_mon_prot_cis.label_asym_id`

A component of the identifier for the monomer. This data item is a pointer to `_atom_site.label_asym_id` in the ATOM\_SITE category.

\* `_struct_mon_prot_cis.label_comp_id`

A component of the identifier for the monomer. This data item is a pointer to `_atom_site.label_comp_id` in the ATOM\_SITE category.

\* `_struct_mon_prot_cis.label_seq_id`

A component of the identifier for the monomer. This data item is a pointer to `_atom_site.label_seq_id` in the ATOM\_SITE category.

### STRUCT\_NCS\_DOM

Data items in the STRUCT\_NCS\_DOM category record information about the domains in an ensemble of domains related by one or more noncrystallographic symmetry operators. A domain need not correspond to a complete polypeptide chain; it can be composed of one or more segments in a single chain, or by segments from more than one chain.

Category group(s): `inclusive_group`  
`struct_group`  
 Category key(s): `_struct_ncs_dom.id`

Example 1 – based on laboratory records for the collagen-like peptide, HYP-.

```
loop_
_struct_ncs_dom.id
_struct_ncs_dom.details
d1 'Chains A, B, and C'
d2 'Chains D, E, and F'
```

`_struct_ncs_dom.details` (text)

A description of special aspects of the structural elements that comprise a domain in an ensemble of domains related by noncrystallographic symmetry.

Example:

```
; The loop between residues 18 and 23 in this domain interacts
with a symmetry-related molecule, and thus deviates
significantly from the noncrystallographic threefold.
```

;

[struct\_ncs\_dom]

\* **struct\_ncs\_dom\_id** (code)  
The value of struct\_ncs\_dom\_id must uniquely identify a record in the STRUCT\_NCS\_DOM list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

```
struct_ncs_dom_lim.dom_id,
struct_ncs_ens_gen.dom_id_1,
struct_ncs_ens_gen.dom_id_2.           [struct_ncs_dom]
```

### STRUCT\_NCS\_DOM\_LIM

Data items in the STRUCT\_NCS\_DOM\_LIM category identify the start and end points of polypeptide chain segments that form all or part of a domain in an ensemble of domains related by non-crystallographic symmetry.

Category group(s): inclusive\_group

struct\_group

Category key(s): struct\_ncs\_dom\_lim.dom\_id  
struct\_ncs\_dom\_lim.beg\_label\_alt\_id  
struct\_ncs\_dom\_lim.beg\_label\_asym\_id  
struct\_ncs\_dom\_lim.beg\_label\_comp\_id  
struct\_ncs\_dom\_lim.beg\_label\_seq\_id  
struct\_ncs\_dom\_lim.end\_label\_alt\_id  
struct\_ncs\_dom\_lim.end\_label\_asym\_id  
struct\_ncs\_dom\_lim.end\_label\_comp\_id  
struct\_ncs\_dom\_lim.end\_label\_seq\_id

Example 1 – based on laboratory records for the collagen-like peptide, HYP-.

```
loop_
  struct_ncs_dom_lim.dom_id
  struct_ncs_dom_lim.beg_label_alt_id
  struct_ncs_dom_lim.beg_label_asym_id
  struct_ncs_dom_lim.beg_label_comp_id
  struct_ncs_dom_lim.beg_label_seq_id
  struct_ncs_dom_lim.end_label_alt_id
  struct_ncs_dom_lim.end_label_asym_id
  struct_ncs_dom_lim.end_label_comp_id
  struct_ncs_dom_lim.end_label_seq_id
  d1 . A PRO 1 . A GLY 29
  d1 . B PRO 31 . B GLY 59
  d1 . C PRO 61 . B GLY 89
  d2 . D PRO 91 . D GLY 119
  d2 . E PRO 121 . E GLY 149
  d2 . F PRO 151 . F GLY 179
```

**struct\_ncs\_dom\_lim.beg\_auth\_asym\_id**  
A component of the identifier for the monomer at which this segment of the domain begins. This data item is a pointer to atom\_site.auth\_asym\_id in the ATOM\_SITE category.

**struct\_ncs\_dom\_lim.beg\_auth\_comp\_id**  
A component of the identifier for the monomer at which this segment of the domain begins. This data item is a pointer to atom\_site.auth\_comp\_id in the ATOM\_SITE category.

**struct\_ncs\_dom\_lim.beg\_auth\_seq\_id**  
A component of the identifier for the monomer at which this segment of the domain begins. This data item is a pointer to atom\_site.auth\_seq\_id in the ATOM\_SITE category.

\* **struct\_ncs\_dom\_lim.beg\_label\_alt\_id**  
A component of the identifier for the monomer at which this segment of the domain begins. This data item is a pointer to atom\_sites\_alt.id in the ATOM\_SITES\_ALT category.

\* **struct\_ncs\_dom\_lim.beg\_label\_asym\_id**  
A component of the identifier for the monomer at which this segment of the domain begins. This data item is a pointer to atom\_site.label\_asym\_id in the ATOM\_SITE category.

\* **struct\_ncs\_dom\_lim.beg\_label\_comp\_id**  
A component of the identifier for the monomer at which this segment of the domain begins. This data item is a pointer to atom\_site.label\_comp\_id in the ATOM\_SITE category.

\* **struct\_ncs\_dom\_lim.beg\_label\_seq\_id**  
A component of the identifier for the monomer at which this segment of the domain begins. This data item is a pointer to atom\_site.label\_seq\_id in the ATOM\_SITE category.

\* **struct\_ncs\_dom\_lim.dom\_id**  
This data item is a pointer to struct\_ncs\_dom\_id in the STRUCT\_NCS\_DOM category.

**struct\_ncs\_dom\_lim.end\_auth\_asym\_id**  
A component of the identifier for the monomer at which this segment of the domain ends. This data item is a pointer to atom\_site.auth\_asym\_id in the ATOM\_SITE category.

**struct\_ncs\_dom\_lim.end\_auth\_comp\_id**  
A component of the identifier for the monomer at which this segment of the domain ends. This data item is a pointer to atom\_site.auth\_comp\_id in the ATOM\_SITE category.

**struct\_ncs\_dom\_lim.end\_auth\_seq\_id**  
A component of the identifier for the monomer at which this segment of the domain ends. This data item is a pointer to atom\_site.auth\_seq\_id in the ATOM\_SITE category.

\* **struct\_ncs\_dom\_lim.end\_label\_alt\_id**  
A component of the identifier for the monomer at which this segment of the domain ends. This data item is a pointer to atom\_sites\_alt.id in the ATOM\_SITES\_ALT category.

\* **struct\_ncs\_dom\_lim.end\_label\_asym\_id**  
A component of the identifier for the monomer at which this segment of the domain ends. This data item is a pointer to atom\_site.label\_asym\_id in the ATOM\_SITE category.

\* **struct\_ncs\_dom\_lim.end\_label\_comp\_id**  
A component of the identifier for the monomer at which this segment of the domain ends. This data item is a pointer to atom\_site.label\_comp\_id in the ATOM\_SITE category.

\* **struct\_ncs\_dom\_lim.end\_label\_seq\_id**  
A component of the identifier for the monomer at which this segment of the domain ends. This data item is a pointer to atom\_site.label\_seq\_id in the ATOM\_SITE category.

### STRUCT\_NCS\_ENS

Data items in the STRUCT\_NCS\_ENS category record information about ensembles of domains related by noncrystallographic symmetry. The point group of the ensemble when taken as a whole may be specified, as well as any special aspects of the ensemble that require description.

Category group(s): inclusive\_group

struct\_group

Category key(s): struct\_ncs\_ens.id

Example 1 – based on laboratory records for the collagen-like peptide, HYP-.

```
struct_ncs_ens.id           en1
struct_ncs_ens.details
; The ensemble represents the pseudo-twofold symmetry
  between domains d1 and d2.
;
```

**\_struct\_ncs\_ens.details** (text)

A description of special aspects of the ensemble.

Example:

```
; The ensemble has a slight translation between domains 1
and 4, but overall it can accurately be described as point
group 222
```

```
; [struct_ncs_ens]
```

\* **\_struct\_ncs\_ens.id** (code)

The value of `_struct_ncs_ens.id` must uniquely identify a record in the STRUCT\_NCS\_ENS list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

```
_struct_ncs_ens_gen.ens_id. [struct_ncs_ens]
```

**\_struct\_ncs\_ens.point\_group** (line)

The point group of the ensemble of structural elements related by one or more noncrystallographic symmetry operations. The relationships need not be precise; this data item is intended to give a rough description of the noncrystallographic symmetry relationships.

Examples: '3', '422', 'non-proper'. [struct\_ncs\_ens]

### STRUCT\_NCS\_ENS\_GEN

Data items in the STRUCT\_NCS\_ENS\_GEN category list domains related by a noncrystallographic symmetry operation and identify the operator.

Category group(s): `inclusive_group`  
`struct_group`

Category key(s): `_struct_ncs_ens_gen.ens_id`  
`_struct_ncs_ens_gen.dom_id_1`  
`_struct_ncs_ens_gen.dom_id_2`  
`_struct_ncs_ens_gen.oper_id`

Example 1 – based on laboratory records for the collagen-like peptide, HYP-.

```
_struct_ncs_ens_gen.dom_id_1 d1
_struct_ncs_ens_gen.dom_id_2 d2
_struct_ncs_ens_gen.ens_id en1
_struct_ncs_ens_gen.oper_id ncsop1
```

\* **\_struct\_ncs\_ens\_gen.dom\_id\_1**

The identifier for the domain that will remain unchanged by the transformation operator. This data item is a pointer to `_struct_ncs_dom.id` in the STRUCT\_NCS\_DOM category.

\* **\_struct\_ncs\_ens\_gen.dom\_id\_2**

The identifier for the domain that will be transformed by application of the transformation operator. This data item is a pointer to `_struct_ncs_dom.id` in the STRUCT\_NCS\_DOM category.

\* **\_struct\_ncs\_ens\_gen.ens\_id**

This data item is a pointer to `_struct_ncs_ens.id` in the STRUCT\_NCS\_ENS category.

\* **\_struct\_ncs\_ens\_gen.oper\_id**

This data item is a pointer to `_struct_ncs_oper.id` in the STRUCT\_NCS\_OPER category.

### STRUCT\_NCS\_OPER

Data items in the STRUCT\_NCS\_OPER category describe the non-crystallographic symmetry operations. Each operator is specified as a matrix and a subsequent translation vector. Operators need not represent proper rotations.

Category group(s): `inclusive_group`  
`struct_group`

Category key(s): `_struct_ncs_oper.id`

Example 1 – based on laboratory records for the protein NS1.

```
_struct_ncs_oper.id ncsop1
_struct_ncs_oper.code given
_struct_ncs_oper.matrix[1][1] 0.247
_struct_ncs_oper.matrix[1][2] 0.935
_struct_ncs_oper.matrix[1][3] 0.256
_struct_ncs_oper.matrix[2][1] 0.929
_struct_ncs_oper.matrix[2][2] 0.153
_struct_ncs_oper.matrix[2][3] 0.337
_struct_ncs_oper.matrix[3][1] 0.276
_struct_ncs_oper.matrix[3][2] 0.321
_struct_ncs_oper.matrix[3][3] -0.906
_struct_ncs_oper.vector[1] -8.253
_struct_ncs_oper.vector[2] -11.743
_struct_ncs_oper.vector[3] -1.782
_struct_ncs_oper.details
; Matrix and translation vector for pseudo-twofold operation.
;
```

**\_struct\_ncs\_oper.code** (code)

A code to indicate whether this operator describes a relationship between coordinates all of which are given in the data block (in which case the value of code is 'given'), or whether the operator is used to generate new coordinates from those that are given in the data block (in which case the value of code is 'generate').

The data value must be one of the following:

<code>given</code>	operator relates coordinates given in the data block
<code>generate</code>	operator generates new coordinates from those given in the data block

[struct\_ncs\_oper]

**\_struct\_ncs\_oper.details** (text)

A description of special aspects of the noncrystallographic symmetry operator.

Example:

```
; The operation is given as a precise threefold rotation,
despite the fact the best rms fit between domain 1 and domain
2 yields a rotation of 119.7 degrees and a translation of
0.13 angstroms.
```

```
; [struct_ncs_oper]
```

\* **\_struct\_ncs\_oper.id** (code)

The value of `_struct_ncs_oper.id` must uniquely identify a record in the STRUCT\_NCS\_OPER list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

```
_struct_ncs_ens_gen.oper_id. [struct_ncs_oper]
```

**\_struct\_ncs\_oper.matrix[1][1]** (float)

The [1][1] element of the  $3 \times 3$  matrix component of a noncrystallographic symmetry operation.

[struct\_ncs\_oper]

**\_struct\_ncs\_oper.matrix[1][2]** (float)

The [1][2] element of the  $3 \times 3$  matrix component of a noncrystallographic symmetry operation.

[struct\_ncs\_oper]

**\_struct\_ncs\_oper.matrix[1][3]** (float)

The [1][3] element of the  $3 \times 3$  matrix component of a noncrystallographic symmetry operation.

[struct\_ncs\_oper]

**\_struct\_ncs\_oper.matrix[2][1]** (float)  
The [2][1] element of the 3 × 3 matrix component of a noncrystallographic symmetry operation.

[struct\_ncs\_oper]

**\_struct\_ncs\_oper.matrix[2][2]** (float)  
The [2][2] element of the 3 × 3 matrix component of a noncrystallographic symmetry operation.

[struct\_ncs\_oper]

**\_struct\_ncs\_oper.matrix[2][3]** (float)  
The [2][3] element of the 3 × 3 matrix component of a noncrystallographic symmetry operation.

[struct\_ncs\_oper]

**\_struct\_ncs\_oper.matrix[3][1]** (float)  
The [3][1] element of the 3 × 3 matrix component of a noncrystallographic symmetry operation.

[struct\_ncs\_oper]

**\_struct\_ncs\_oper.matrix[3][2]** (float)  
The [3][2] element of the 3 × 3 matrix component of a noncrystallographic symmetry operation.

[struct\_ncs\_oper]

**\_struct\_ncs\_oper.matrix[3][3]** (float)  
The [3][3] element of the 3 × 3 matrix component of a noncrystallographic symmetry operation.

[struct\_ncs\_oper]

**\_struct\_ncs\_oper.vector[1]** (float)  
The [1] element of the three-element vector component of a noncrystallographic symmetry operation.

[struct\_ncs\_oper]

**\_struct\_ncs\_oper.vector[2]** (float)  
The [2] element of the three-element vector component of a noncrystallographic symmetry operation.

[struct\_ncs\_oper]

**\_struct\_ncs\_oper.vector[3]** (float)  
The [3] element of the three-element vector component of a noncrystallographic symmetry operation.

[struct\_ncs\_oper]

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
  _struct_ref.id
  _struct_ref.entity_id
  _struct_ref.biol_id
  _struct_ref.db_name
  _struct_ref.db_code
  _struct_ref.seq_align
  _struct_ref.seq_dif
  _struct_ref.details
  seq_pdb . 1 PDB 5HVP .
; The structure of the closely related compound,
  isobutyryl-pepstatin (pepstatin A) in complex with
  rhizopuspepsin
;
  seq_genbank 1 . GenBank AAG30358 complete yes .
```

**\_struct\_ref.biol\_id**  
This data item is a pointer to **\_struct\_biol.id** in the STRUCT\_BIOL category.

\* **\_struct\_ref.db\_code** (line)  
The code for this entity or biological unit or for a closely related entity or biological unit in the named database.  
Examples: '1ABC', 'ABCDEF'. [struct\_ref]

\* **\_struct\_ref.db\_name** (line)  
The name of the database containing reference information about this entity or biological unit.  
Examples: 'PDB', 'CSD', 'Genbank'. [struct\_ref]

**\_struct\_ref.details** (text)  
A description of special aspects of the relationship between the entity or biological unit described in the data block and that in the referenced database entry. [struct\_ref]

\* **\_struct\_ref.entity\_id**  
This data item is a pointer to **\_entity.id** in the ENTITY category.

\* **\_struct\_ref.id** (code)  
The value of **\_struct\_ref.id** must uniquely identify a record in the STRUCT\_REF list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

**\_struct\_ref.seq\_ref\_id**. [struct\_ref]

**\_struct\_ref.seq\_align** (ucode)  
A flag to indicate the scope of the alignment between the sequence of the entity or biological unit described in the data block and that in the referenced database entry. 'complete' indicates that alignment spans the entire length of both sequences (although point differences may occur and can be annotated using the data items in the STRUCT\_REF\_SEQ\_DIF category). 'partial' indicates a partial alignment. The region (or regions) of the alignment may be delimited using data items in the STRUCT\_REF\_SEQ category. This data item may also take the value '.', indicating that the reference is not to a sequence.

The data value must be one of the following:

complete	alignment is complete
partial	alignment is partial
.	reference is not to a sequence

[struct\_ref]

### STRUCT\_REF

Data items in the STRUCT\_REF category allow the author of a data block to relate the entities or biological units described in the data block to information archived in external databases. For references to the sequence of a polymer, the value of the data item **\_struct\_ref.seq\_align** is used to indicate whether the correspondence between the sequence of the entity or biological unit in the data block and the sequence in the referenced database entry is 'complete' or 'partial'. If this value is 'partial', the region (or regions) of the alignment may be delimited using data items in the STRUCT\_REF\_SEQ category. Similarly, the value of **\_struct\_ref.seq\_dif** is used to indicate whether the two sequences contain point differences. If the value is 'yes', the differences may be identified and annotated using data items in the STRUCT\_REF\_SEQ\_DIF category.

Category group(s): **inclusive\_group**

**struct\_group**

Category key(s): **\_struct\_ref.id**

**\_struct\_ref.seq\_dif** (ucode)  
 A flag to indicate the presence ('yes') or absence ('no') of point differences between the sequence of the entity or biological unit described in the data block and that in the referenced database entry. This data item may also take the value '.', indicating that the reference is not to a sequence.

The data value must be one of the following:

no	there are no point differences
n	abbreviation for 'no'
yes	there are point difference
y	abbreviation for 'yes'
.	reference is not to a sequence

[struct\_ref]

### STRUCT\_REF\_SEQ

Data items in the STRUCT\_REF\_SEQ category provide a mechanism for indicating and annotating a region (or regions) of alignment between the sequence of an entity or biological unit described in the data block and the sequence in the referenced database entry.

Category group(s): **inclusive\_group**  
**struct\_group**

Category key(s): **\_struct\_ref\_seq.align\_id**

*Example 1 – based on the sequence alignment of CHER from M. xantus (36 to 288) and CHER from S. typhimurium (18 to 276).*

```
_struct_ref_seq.align_id      alg1
_struct_ref_seq.ref_id       seqdb1
_struct_ref_seq.seq_align_beg 36
_struct_ref_seq.seq_align_end 288
_struct_ref_seq.db_align_beg  18
_struct_ref_seq.db_align_end  276
_struct_ref_seq.details
; The alignment contains 3 gaps larger than 2 residues
;
```

\* **\_struct\_ref\_seq.align\_id** (code)

The value of **\_struct\_ref\_seq.align\_id** must uniquely identify a record in the STRUCT\_REF\_SEQ list. Note that this item need not be a number; it can be any unique identifier.

*The following item(s) have an equivalent role in their respective categories:*

**\_struct\_ref\_seq.dif.align\_id** [struct\_ref\_seq]

\* **\_struct\_ref\_seq.db\_align\_beg** (int)

The sequence position in the referenced database entry at which the alignment begins.

\* **\_struct\_ref\_seq.db\_align\_end** (int)

The sequence position in the referenced database entry at which the alignment ends.

**\_struct\_ref\_seq.details** (text)

A description of special aspects of the sequence alignment.

[struct\_ref\_seq]

\* **\_struct\_ref\_seq.ref\_id**

This data item is a pointer to **\_struct\_ref.id** in the STRUCT\_REF category.

\* **\_struct\_ref\_seq.seq\_align\_beg**

The sequence position in the entity or biological unit described in the data block at which the alignment begins. This data item is a pointer to **\_entity\_poly\_seq.num** in the ENTITY\_POLY\_SEQ category.

\* **\_struct\_ref\_seq.seq\_align\_end**

The sequence position in the entity or biological unit described in the data block at which the alignment ends. This data item is a pointer to **\_entity\_poly\_seq.num** in the ENTITY\_POLY\_SEQ category.

### STRUCT\_REF\_SEQ\_DIF

Data items in the STRUCT\_REF\_SEQ\_DIF category provide a mechanism for indicating and annotating point differences between the sequence of the entity or biological unit described in the data block and the sequence of the referenced database entry.

Category group(s): **inclusive\_group**  
**struct\_group**

Category key(s): **\_struct\_ref\_seq\_dif.align\_id**  
**\_struct\_ref\_seq\_dif.seq\_num**

*Example 1 – based on laboratory records for CAP-DNA complex.*

```
_struct_ref_seq_dif.align_id  align2
_struct_ref_seq_dif.seq_num   181
_struct_ref_seq_dif.db_mon_id  GLU
_struct_ref_seq_dif.mon_id    PHE
_struct_ref_seq_dif.details
; A point mutation was introduced in the CAP at position 181
; substituting PHE for GLU.
;
```

\* **\_struct\_ref\_seq\_dif.align\_id**

This data item is a pointer to **\_struct\_ref\_seq.align\_id** in the STRUCT\_REF\_SEQ category.

\* **\_struct\_ref\_seq\_dif.db\_mon\_id**

The monomer type found at this position in the referenced database entry. This data item is a pointer to **\_chem\_comp.id** in the CHEM\_COMP category.

**\_struct\_ref\_seq\_dif.details** (text)

A description of special aspects of the point differences between the sequence of the entity or biological unit described in the data block and that in the referenced database entry.

[struct\_ref\_seq\_dif]

\* **\_struct\_ref\_seq\_dif.mon\_id**

The monomer type found at this position in the sequence of the entity or biological unit described in this data block. This data item is a pointer to **\_chem\_comp.id** in the CHEM\_COMP category.

\* **\_struct\_ref\_seq\_dif.seq\_num**

This data item is a pointer to **\_entity\_poly\_seq.num** in the ENTITY\_POLY\_SEQ category.

**STRUCT\_SHEET**

Data items in the STRUCT\_SHEET category record details about the  $\beta$ -sheets.

Category group(s): **inclusive\_group**  
**struct\_group**  
 Category key(s): **\_struct\_sheet.id**

---

*Example 1 - simple beta-barrel.*

```

N O   N O   N O   N O   N O   N O
10--11--12--13--14--15--16--17--18--19--20 strand_a
  N O   N O   N O   N O   N O   N O
  / \   / \   / \   / \   / \
N O   N O   N O   N O   N O   N O
30--31--32--33--34--35--36--37--38--39--40 strand_b
  N O   N O   N O   N O   N O   N O
  / \   / \   / \   / \   / \
N O   N O   N O   N O   N O   N O
50--51--52--53--54--55--56--57--58--59--60 strand_c
  N O   N O   N O   N O   N O   N O
  / \   / \   / \   / \   / \
N O   N O   N O   N O   N O   N O
70--71--72--73--74--75--76--77--78--79--80 strand_d
  N O   N O   N O   N O   N O   N O
  / \   / \   / \   / \   / \
N O   N O   N O   N O   N O   N O
90--91--92--93--94--95--96--97--98--99--100 strand_e
  N O   N O   N O   N O   N O   N O
  / \   / \   / \   / \   / \
N O   N O   N O   N O   N O   N O
110-111-112-113-114-115-116-117-118-119-120 strand_f
  N O   N O   N O   N O   N O   N O
  / \   / \   / \   / \   / \
N O   N O   N O   N O   N O   N O
130-131-132-133-134-135-136-137-138-139-140 strand_g
  N O   N O   N O   N O   N O   N O
  / \   / \   / \   / \   / \
N O   N O   N O   N O   N O   N O
150-151-152-153-154-155-156-157-158-159-160 strand_h
  N O   N O   N O   N O   N O   N O
  / \   / \   / \   / \   / \
  
```

```

_struct_sheet.id      sheet_1
_struct_sheet.type    'beta-barrel'
_struct_sheet.number_strands 8
_struct_sheet.details .
  
```

---

*Example 2 - five stranded mixed-sense sheet with one two-piece strand.*

```

      N O   N O   N O   N O
-10--11--12--13--14--15--16--17--18-> strand_a
N O   N O   N O   N O   N O
| |   | |   | |   | |
O N   O N   O N   O N   O N
<-119-118-117-116-115-114-113-112-111-110- strand_b
  O N   O N   O N   O N   O N
  \ /   \ /   \ /   \ /   \ /
  O N   O N   O N   O N   O N   O N
      <-41--40--39--38--37--36--35--34--33--32--31--30- strand_c
        O N   O N   O N   O N   O N   O N
        | |   | |   | |   | |
        N O   N O   N O   N O   N O   N O
strand_d1 -50--51--52->   -90--91--92--93--95--95--96--97-> strand_d2
        N O   N O   N O   N O   N O
        | |   | |   | |   | |
        O N   O N   O N   O N   O N   O N
      <-80--79--78--77--76--75--74--73--72--71--70- strand_e
        O N   O N   O N   O N   O N
  
```

```

_struct_sheet.id      sheet_2
_struct_sheet.type    'five stranded, mixed-sense'
_struct_sheet.number_strands 5
_struct_sheet.details 'strand_d is in two pieces'
  
```

**\_struct\_sheet.details**

(text)

A description of special aspects of the  $\beta$ -sheet.

[struct\_sheet]

**\*\_struct\_sheet.id**

(code)

The value of `_struct_sheet.id` must uniquely identify a record in the STRUCT\_SHEET list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

`_struct_sheet_hbond.sheet_id,``_struct_sheet_order.sheet_id,``_struct_sheet_range.sheet_id,``_struct_sheet_topology.sheet_id.`

[struct\_sheet]

**\_struct\_sheet.number\_strands**

(int)

The number of strands in the sheet. If a given range of residues bulges out from the strands, it is still counted as one strand. If a strand is composed of two different regions of polypeptide, it is still counted as one strand, as long as the proper hydrogen-bonding connections are made to adjacent strands.

[struct\_sheet]

**\_struct\_sheet.type**

(text)

A simple descriptor for the type of the sheet.

Examples: 'jelly-roll', 'Rossmann fold', 'beta barrel'. [struct\_sheet]

**STRUCT\_SHEET\_HBOND**

Data items in the STRUCT\_SHEET\_HBOND category record details about the hydrogen bonding between residue ranges in a  $\beta$ -sheet. It is necessary to treat hydrogen bonding independently of the designation of ranges, because the hydrogen bonding may begin in different places for the interactions of a given strand with the one preceding it and the one following it in the sheet.

Category group(s): `inclusive_group`

`struct_group`

Category key(s): `_struct_sheet_hbond.sheet_id`  
`_struct_sheet_hbond.range_id_1`  
`_struct_sheet_hbond.range_id_2`

Example 1 – simple beta-barrel.

```

loop_
_struct_sheet_hbond.sheet_id
_struct_sheet_hbond.range_id_1
_struct_sheet_hbond.range_id_2
_struct_sheet_hbond.range_1_beg_label_seq_id
_struct_sheet_hbond.range_1_beg_label_atom_id
_struct_sheet_hbond.range_2_beg_label_seq_id
_struct_sheet_hbond.range_2_beg_label_atom_id
_struct_sheet_hbond.range_1_end_label_seq_id
_struct_sheet_hbond.range_1_end_label_atom_id
_struct_sheet_hbond.range_2_end_label_seq_id
_struct_sheet_hbond.range_2_end_label_atom_id
sheet_1 strand_a strand_b 11 N 30 O 19 O 40 N
sheet_1 strand_b strand_c 31 N 50 O 39 O 60 N
sheet_1 strand_c strand_d 51 N 70 O 59 O 80 N
sheet_1 strand_d strand_e 71 N 90 O 89 O 100 N
sheet_1 strand_e strand_f 91 N 110 O 99 O 120 N
sheet_1 strand_f strand_g 111 N 130 O 119 O 140 N
sheet_1 strand_g strand_h 131 N 150 O 139 O 160 N
sheet_1 strand_h strand_a 151 N 10 O 159 O 180 N

```

Example 2 – five stranded mixed-sense sheet with one two-piece strand.

```

loop_
_struct_sheet_hbond.sheet_id
_struct_sheet_hbond.range_id_1
_struct_sheet_hbond.range_id_2
_struct_sheet_hbond.range_1_beg_label_seq_id
_struct_sheet_hbond.range_1_beg_label_atom_id
_struct_sheet_hbond.range_2_beg_label_seq_id
_struct_sheet_hbond.range_2_beg_label_atom_id
_struct_sheet_hbond.range_1_end_label_seq_id
_struct_sheet_hbond.range_1_end_label_atom_id
_struct_sheet_hbond.range_2_end_label_seq_id
_struct_sheet_hbond.range_2_end_label_atom_id
sheet_2 strand_a strand_b 20 N 119 O 18 O 111 N
sheet_2 strand_b strand_c 110 N 33 O 118 N 41 O
sheet_2 strand_c strand_d1 38 N 52 O 40 O 50 N
sheet_2 strand_c strand_d2 30 N 96 O 36 O 90 N
sheet_2 strand_d1 strand_e 51 N 80 O 51 O 80 N
sheet_2 strand_d2 strand_e 91 N 76 O 97 O 70 N

```

**\_struct\_sheet\_hbond.range\_1\_beg\_auth\_atom\_id**

A component of the identifier for the residue for the first partner of the first hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.auth_atom_id` in the ATOM\_SITE category.

**\_struct\_sheet\_hbond.range\_1\_beg\_auth\_seq\_id**

A component of the identifier for the residue for the first partner of the first hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.auth_seq_id` in the ATOM\_SITE category.

**\*\_struct\_sheet\_hbond.range\_1\_beg\_label\_atom\_id**

A component of the identifier for the residue for the first partner of the first hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.label_atom_id` in the ATOM\_SITE category.

**\*\_struct\_sheet\_hbond.range\_1\_beg\_label\_seq\_id**

A component of the identifier for the residue for the first partner of the first hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.label_seq_id` in the ATOM\_SITE category.

**\_struct\_sheet\_hbond.range\_1\_end\_auth\_atom\_id**

A component of the identifier for the residue for the first partner of the last hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.auth_atom_id` in the ATOM\_SITE category.

**\_struct\_sheet\_hbond.range\_1\_end\_auth\_seq\_id**

A component of the identifier for the residue for the first partner of the last hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.auth_seq_id` in the ATOM\_SITE category.

**\*\_struct\_sheet\_hbond.range\_1\_end\_label\_atom\_id**

A component of the identifier for the residue for the first partner of the last hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.label_atom_id` in the ATOM\_SITE category.

**\*\_struct\_sheet\_hbond.range\_1\_end\_label\_seq\_id**

A component of the identifier for the residue for the first partner of the last hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.label_seq_id` in the ATOM\_SITE category.

**\_struct\_sheet\_hbond.range\_2\_beg\_auth\_atom\_id**

A component of the identifier for the residue for the second partner of the first hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.auth_atom_id` in the `ATOM_SITE` category.

**\_struct\_sheet\_hbond.range\_2\_beg\_auth\_seq\_id**

A component of the identifier for the residue for the second partner of the first hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.auth_seq_id` in the `ATOM_SITE` category.

**\*\_struct\_sheet\_hbond.range\_2\_beg\_label\_atom\_id**

A component of the identifier for the residue for the second partner of the first hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.label_atom_id` in the `ATOM_SITE` category.

**\*\_struct\_sheet\_hbond.range\_2\_beg\_label\_seq\_id**

A component of the identifier for the residue for the second partner of the first hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.label_seq_id` in the `ATOM_SITE` category.

**\_struct\_sheet\_hbond.range\_2\_end\_auth\_atom\_id**

A component of the identifier for the residue for the second partner of the last hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.auth_atom_id` in the `ATOM_SITE` category.

**\_struct\_sheet\_hbond.range\_2\_end\_auth\_seq\_id**

A component of the identifier for the residue for the second partner of the last hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.auth_seq_id` in the `ATOM_SITE` category.

**\*\_struct\_sheet\_hbond.range\_2\_end\_label\_atom\_id**

A component of the identifier for the residue for the second partner of the last hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.label_atom_id` in the `ATOM_SITE` category.

**\*\_struct\_sheet\_hbond.range\_2\_end\_label\_seq\_id**

A component of the identifier for the residue for the second partner of the last hydrogen bond between two residue ranges in a sheet. This data item is a pointer to `_atom_site.label_seq_id` in the `ATOM_SITE` category.

**\*\_struct\_sheet\_hbond.range\_id\_1**

This data item is a pointer to `_struct_sheet_range.id` in the `STRUCT_SHEET_RANGE` category.

**\*\_struct\_sheet\_hbond.range\_id\_2**

This data item is a pointer to `_struct_sheet_range.id` in the `STRUCT_SHEET_RANGE` category.

**\*\_struct\_sheet\_hbond.sheet\_id**

This data item is a pointer to `_struct_sheet.id` in the `STRUCT_SHEET` category.

**STRUCT\_SHEET\_ORDER**

Data items in the `STRUCT_SHEET_ORDER` category record details about the order of the residue ranges that form a  $\beta$ -sheet. All order links are pairwise and the specified pairs are assumed to be adjacent to one another in the sheet. These data items are an alternative to the `STRUCT_SHEET_TOPOLOGY` data items and they allow all manner of sheets to be described.

Category group(s): `inclusive_group`  
`struct_group`

Category key(s): `_struct_sheet_order.sheet_id`  
`_struct_sheet_order.range_id_1`  
`_struct_sheet_order.range_id_2`

*Example 1 – simple beta-barrel.*

```
loop_
  _struct_sheet_order.sheet_id
  _struct_sheet_order.range_id_1
  _struct_sheet_order.range_id_2
  _struct_sheet_order.offset
  _struct_sheet_order.sense
sheet_1 strand_a strand_b +1 parallel
sheet_1 strand_b strand_c +1 parallel
sheet_1 strand_c strand_d +1 parallel
sheet_1 strand_d strand_e +1 parallel
sheet_1 strand_e strand_f +1 parallel
sheet_1 strand_f strand_g +1 parallel
sheet_1 strand_g strand_h +1 parallel
sheet_1 strand_h strand_a +1 parallel
```

*Example 2 – five stranded mixed-sense sheet with one two-piece strand.*

```
loop_
  _struct_sheet_order.sheet_id
  _struct_sheet_order.range_id_1
  _struct_sheet_order.range_id_2
  _struct_sheet_order.offset
  _struct_sheet_order.sense
sheet_2 strand_a strand_b +1 anti-parallel
sheet_2 strand_b strand_c +1 parallel
sheet_2 strand_c strand_d1 +1 anti-parallel
sheet_2 strand_c strand_d2 +1 anti-parallel
sheet_2 strand_d1 strand_e +1 anti-parallel
sheet_2 strand_d2 strand_e +1 anti-parallel
```

**\_struct\_sheet\_order.offset**

(int)

Designates the relative position in the sheet, plus or minus, of the second residue range to the first.

[struct\_sheet\_order]

**\*\_struct\_sheet\_order.range\_id\_1**

This data item is a pointer to `_struct_sheet_range.id` in the `STRUCT_SHEET_RANGE` category.

**\*\_struct\_sheet\_order.range\_id\_2**

This data item is a pointer to `_struct_sheet_range.id` in the `STRUCT_SHEET_RANGE` category.

**\_struct\_sheet\_order.sense**

(ucode)

A flag to indicate whether the two designated residue ranges are parallel or antiparallel to one another.

The data value must be one of the following:

```
parallel
anti-parallel
```

[struct\_sheet\_order]

**\*\_struct\_sheet\_order.sheet\_id**

This data item is a pointer to `_struct_sheet.id` in the `STRUCT_SHEET` category.

**STRUCT\_SHEET\_RANGE**

Data items in the STRUCT\_SHEET\_RANGE category record details about the residue ranges that form a  $\beta$ -sheet. Residues are included in a range if they made  $\beta$ -sheet-type hydrogen-bonding interactions with at least one adjacent strand and if there are at least two residues in the range.

Category group(s): inclusive\_group

struct\_group

Category key(s): \_struct\_sheet\_range.sheet\_id

\_struct\_sheet\_range.id

*Example 1 – simple beta-barrel.*

```
loop_
_struct_sheet_range.sheet_id
_struct_sheet_range.id
_struct_sheet_range.beg_label_comp_id
_struct_sheet_range.beg_label_asym_id
_struct_sheet_range.beg_label_seq_id
_struct_sheet_range.end_label_comp_id
_struct_sheet_range.end_label_asym_id
_struct_sheet_range.end_label_seq_id
_struct_sheet_range.symmetry
sheet_1 strand_a ala A 20 ala A 30 1_555
sheet_1 strand_b ala A 40 ala A 50 1_555
sheet_1 strand_c ala A 60 ala A 70 1_555
sheet_1 strand_d ala A 80 ala A 90 1_555
sheet_1 strand_e ala A 100 ala A 110 1_555
sheet_1 strand_f ala A 120 ala A 130 1_555
sheet_1 strand_g ala A 140 ala A 150 1_555
sheet_1 strand_h ala A 160 ala A 170 1_555
```

*Example 2 – five stranded mixed-sense sheet with one two-piece strand.*

```
loop_
_struct_sheet_range.sheet_id
_struct_sheet_range.id
_struct_sheet_range.beg_label_comp_id
_struct_sheet_range.beg_label_asym_id
_struct_sheet_range.beg_label_seq_id
_struct_sheet_range.end_label_comp_id
_struct_sheet_range.end_label_asym_id
_struct_sheet_range.end_label_seq_id
_struct_sheet_range.symmetry
sheet_2 strand_a ala A 10 ala A 18 1_555
sheet_2 strand_b ala A 110 ala A 119 1_555
sheet_2 strand_c ala A 30 ala A 41 1_555
sheet_2 strand_d1 ala A 50 ala A 52 1_555
sheet_2 strand_d2 ala A 90 ala A 97 1_555
sheet_2 strand_e ala A 70 ala A 80 1_555
```

**\_struct\_sheet\_range.beg\_auth\_asym\_id**

A component of the identifier for the residue at which the  $\beta$ -sheet range begins. This data item is a pointer to `_atom_site.auth_asym_id` in the ATOM\_SITE category.

**\_struct\_sheet\_range.beg\_auth\_comp\_id**

A component of the identifier for the residue at which the  $\beta$ -sheet range begins. This data item is a pointer to `_atom_site.auth_comp_id` in the ATOM\_SITE category.

**\_struct\_sheet\_range.beg\_auth\_seq\_id**

A component of the identifier for the residue at which the  $\beta$ -sheet range begins. This data item is a pointer to `_atom_site.auth_seq_id` in the ATOM\_SITE category.

**\*\_struct\_sheet\_range.beg\_label\_asym\_id**

A component of the identifier for the residue at which the  $\beta$ -sheet range begins. This data item is a pointer to `_struct_asym.id` in the STRUCT\_ASYM category.

**\*\_struct\_sheet\_range.beg\_label\_comp\_id**

A component of the identifier for the residue at which the  $\beta$ -sheet range begins. This data item is a pointer to `_chem_comp.id` in the CHEM\_COMP category.

**\*\_struct\_sheet\_range.beg\_label\_seq\_id**

A component of the identifier for the residue at which the  $\beta$ -sheet range begins. This data item is a pointer to `_atom_site.label_seq_id` in the ATOM\_SITE category.

**\_struct\_sheet\_range.end\_auth\_asym\_id**

A component of the identifier for the residue at which the  $\beta$ -sheet range ends. This data item is a pointer to `_atom_site.auth_asym_id` in the ATOM\_SITE category.

**\_struct\_sheet\_range.end\_auth\_comp\_id**

A component of the identifier for the residue at which the  $\beta$ -sheet range ends. This data item is a pointer to `_atom_site.auth_comp_id` in the ATOM\_SITE category.

**\_struct\_sheet\_range.end\_auth\_seq\_id**

A component of the identifier for the residue at which the  $\beta$ -sheet range ends. This data item is a pointer to `_atom_site.auth_seq_id` in the ATOM\_SITE category.

**\*\_struct\_sheet\_range.end\_label\_asym\_id**

A component of the identifier for the residue at which the  $\beta$ -sheet range ends. This data item is a pointer to `_struct_asym.id` in the STRUCT\_ASYM category.

**\*\_struct\_sheet\_range.end\_label\_comp\_id**

A component of the identifier for the residue at which the  $\beta$ -sheet range ends. This data item is a pointer to `_chem_comp.id` in the CHEM\_COMP category.

**\*\_struct\_sheet\_range.end\_label\_seq\_id**

A component of the identifier for the residue at which the  $\beta$ -sheet range ends. This data item is a pointer to `_atom_site.label_seq_id` in the ATOM\_SITE category.

**\*\_struct\_sheet\_range.id**

(code)

The value of `_struct_sheet_range.id` must uniquely identify a range in a given sheet in the STRUCT\_SHEET\_RANGE list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

```
_struct_sheet_hbond.range_id_1,
_struct_sheet_hbond.range_id_2,
_struct_sheet_order.range_id_1,
_struct_sheet_order.range_id_2,
_struct_sheet_topology.range_id_1,
_struct_sheet_topology.range_id_2. [struct_sheet_range]
```

**\*\_struct\_sheet\_range.sheet\_id**

This data item is a pointer to `_struct_sheet.id` in the STRUCT\_SHEET category.

**\_struct\_sheet\_range.symmetry**

(symop)

Describes the symmetry operation that should be applied to the residues delimited by the start and end designators in order to generate the appropriate strand in this sheet.

[struct\_sheet\_range]

**STRUCT\_SHEET\_TOPOLOGY**

Data items in the STRUCT\_SHEET\_TOPOLOGY category record details about the topology of the residue ranges that form a  $\beta$ -sheet. All topology links are pairwise and the specified pairs are assumed to be successive in the amino-acid sequence. These data items are useful in describing various simple and complex folds, but they become inadequate when the strands in the sheet come from more than one chain. The STRUCT\_SHEET\_ORDER data items can be used to describe single- and multiple-chain-containing sheets.

Category group(s): **inclusive\_group**  
**struct\_group**

Category key(s): **\_struct\_sheet\_topology.sheet\_id**  
**\_struct\_sheet\_topology.range\_id\_1**  
**\_struct\_sheet\_topology.range\_id\_2**

*Example 1 – simple beta-barrel.*

```
loop_
_struct_sheet_topology.sheet_id
_struct_sheet_topology.range_id_1
_struct_sheet_topology.range_id_2
_struct_sheet_topology.offset
_struct_sheet_topology.sense
sheet_1 strand_a strand_b +1 parallel
sheet_1 strand_b strand_c +1 parallel
sheet_1 strand_c strand_d +1 parallel
sheet_1 strand_d strand_e +1 parallel
sheet_1 strand_e strand_f +1 parallel
sheet_1 strand_f strand_g +1 parallel
sheet_1 strand_g strand_h +1 parallel
sheet_1 strand_h strand_a +1 parallel
```

*Example 2 – five stranded mixed-sense sheet with one two-piece strand.*

```
loop_
_struct_sheet_topology.sheet_id
_struct_sheet_topology.range_id_1
_struct_sheet_topology.range_id_2
_struct_sheet_topology.offset
_struct_sheet_topology.sense
sheet_2 strand_a strand_c +2 anti-parallel
sheet_2 strand_c strand_d1 +1 anti-parallel
sheet_2 strand_d1 strand_e +1 anti-parallel
sheet_2 strand_e strand_d2 -1 anti-parallel
sheet_2 strand_d2 strand_b -2 anti-parallel
```

**\_struct\_sheet\_topology.offset** *(int)*  
Designates the relative position in the sheet, plus or minus, of the second residue range to the first.

[struct\_sheet\_topology]

\* **\_struct\_sheet\_topology.range\_id\_1**  
This data item is a pointer to **\_struct\_sheet\_range.id** in the STRUCT\_SHEET\_RANGE category.

\* **\_struct\_sheet\_topology.range\_id\_2**  
This data item is a pointer to **\_struct\_sheet\_range.id** in the STRUCT\_SHEET\_RANGE category.

**\_struct\_sheet\_topology.sense** *(ucode)*  
A flag to indicate whether the two designated residue ranges are parallel or antiparallel to one another.

The data value must be one of the following:

parallel

anti-parallel

[struct\_sheet\_topology]

\* **\_struct\_sheet\_topology.sheet\_id**  
This data item is a pointer to **\_struct\_sheet.id** in the STRUCT\_SHEET category.

**STRUCT\_SITE**

Data items in the STRUCT\_SITE category record details about portions of the structure that contribute to structurally relevant sites (*e.g.* active sites, substrate-binding subsites, metal-coordination sites).

Category group(s): **inclusive\_group**  
**struct\_group**

Category key(s): **\_struct\_site.id**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
_struct_site.id
_struct_site.details
'P2 site C'
; residues with a contact < 3.7 \%%A to an atom in the P2
moiety of the inhibitor in the conformation with
_struct_asym.id = C
;
'P2 site D'
; residues with a contact < 3.7 \%%A to an atom in the P1
moiety of the inhibitor in the conformation with
_struct_asym.id = D
;
```

**\_struct\_site.details** *(text)*

A description of special aspects of the site.

[struct\_site]

\* **\_struct\_site.id** *(line)*

The value of **\_struct\_site.id** must uniquely identify a record in the STRUCT\_SITE list. Note that this item need not be a number; it can be any unique identifier.

The following item(s) have an equivalent role in their respective categories:

**\_struct\_site\_gen.site\_id**,

**\_struct\_site\_keywords.site\_id**,

**\_struct\_site\_view.site\_id**.

[struct\_site]

**STRUCT\_SITE\_GEN**

Data items in the STRUCT\_SITE\_GEN category record details about the generation of portions of the structure that contribute to structurally relevant sites.

Category group(s): **inclusive\_group**  
**struct\_group**

Category key(s): **\_struct\_site\_gen.id**  
**\_struct\_site\_gen.site\_id**

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
_struct_site_gen.id
_struct_site_gen.site_id
_struct_site_gen.label_comp_id
_struct_site_gen.label_asym_id
_struct_site_gen.label_seq_id
_struct_site_gen.symmetry
_struct_site_gen.details
1 1 VAL A 32 1_555 .
2 1 ILE A 47 1_555 .
3 1 VAL A 82 1_555 .
4 1 ILE A 84 1_555 .
5 2 VAL B 232 1_555 .
6 2 ILE B 247 1_555 .
7 2 VAL B 282 1_555 .
8 2 ILE B 284 1_555 .
```

**\_struct\_site\_gen.auth\_asym\_id**

A component of the identifier for participants in the site. This data item is a pointer to **\_atom\_site.auth\_asym\_id** in the ATOM\_SITE category.

**\_struct\_site\_gen.auth\_atom\_id**

A component of the identifier for participants in the site. This data item is a pointer to `_atom_site.auth_atom_id` in the `ATOM_SITE` category.

**\_struct\_site\_gen.auth\_comp\_id**

A component of the identifier for participants in the site. This data item is a pointer to `_atom_site.auth_comp_id` in the `ATOM_SITE` category.

**\_struct\_site\_gen.auth\_seq\_id**

A component of the identifier for participants in the site. This data item is a pointer to `_atom_site.auth_seq_id` in the `ATOM_SITE` category.

**\_struct\_site\_gen.details** (text)

A description of special aspects of the symmetry generation of this portion of the structural site.

Example:

```
; The zinc atom lies on a special position; application of
symmetry elements to generate the insulin hexamer will
generate excess zinc atoms, which must be removed by hand.
; [struct_site_gen]
```

**\*\_struct\_site\_gen.id** (line)

The value of `_struct_site_gen.id` must uniquely identify a record in the `STRUCT_SITE_GEN` list. Note that this item need not be a number; it can be any unique identifier.

[struct\_site\_gen]

**\*\_struct\_site\_gen.label\_alt\_id**

A component of the identifier for participants in the site. This data item is a pointer to `_atom_sites_alt.id` in the `ATOM_SITES_ALT` category.

**\*\_struct\_site\_gen.label\_asym\_id**

A component of the identifier for participants in the site. This data item is a pointer to `_atom_site.label_asym_id` in the `ATOM_SITE` category.

**\*\_struct\_site\_gen.label\_atom\_id**

A component of the identifier for participants in the site. This data item is a pointer to `_chem_comp_atom.atom_id` in the `CHEM_COMP_ATOM` category.

**\*\_struct\_site\_gen.label\_comp\_id**

A component of the identifier for participants in the site. This data item is a pointer to `_atom_site.label_comp_id` in the `ATOM_SITE` category.

**\*\_struct\_site\_gen.label\_seq\_id**

A component of the identifier for participants in the site. This data item is a pointer to `_atom_site.label_seq_id` in the `ATOM_SITE` category.

**\*\_struct\_site\_gen.site\_id**

This data item is a pointer to `_struct_site.id` in the `STRUCT_SITE` category.

**\_struct\_site\_gen.symmetry** (symop)

Describes the symmetry operation that should be applied to the atom set specified by `_struct_site_gen.label*` to generate a portion of the site.

Examples: '.' (no symmetry or translation to site), '4' (4th symmetry operation applied),

```
'7_645' (7th symmetry position: +a on x, -b on y). [struct_site_gen]
```

**STRUCT\_SITE\_KEYWORDS**

Data items in the `STRUCT_SITE_KEYWORDS` category record keywords describing the site.

Category group(s): `inclusive_group`

`struct_group`

Category key(s): `_struct_site_keywords.site_id`

`_struct_site_keywords.text`

Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.

```
loop_
_struct_site_keywords.site_id
_struct_site_keywords.text
'P2 site C' 'binding site'
'P2 site C' 'binding pocket'
'P2 site C' 'P2 site'
'P2 site C' 'P2 pocket'
'P2 site D' 'binding site'
'P2 site D' 'binding pocket'
'P2 site D' 'P2 site'
'P2 site D' 'P2 pocket'
```

**\*\_struct\_site\_keywords.site\_id**

This data item is a pointer to `_struct_site.id` in the `STRUCT_SITE` category.

**\*\_struct\_site\_keywords.text** (text)

Keywords describing this site.

Examples: 'active site', 'binding pocket', 'Ca coordination'.

[struct\_site\_keywords]

**STRUCT\_SITE\_VIEW**

Data items in the `STRUCT_SITE_VIEW` category record details about how to draw and annotate an informative view of the site.

Category group(s): `inclusive_group`

`struct_group`

Category key(s): `_struct_site_view.id`

Example 1 – based on NDB structure GDL001 by Coll, Aymami, Van Der Marel, Van Boom, Rich & Wang [Biochemistry (1989), 28, 310–320].

```
_struct_site_view.id 1
_struct_site_view.rot_matrix[1][1] 0.132
_struct_site_view.rot_matrix[1][2] 0.922
_struct_site_view.rot_matrix[1][3] -0.363
_struct_site_view.rot_matrix[2][1] 0.131
_struct_site_view.rot_matrix[2][2] -0.380
_struct_site_view.rot_matrix[2][3] -0.916
_struct_site_view.rot_matrix[3][1] -0.982
_struct_site_view.rot_matrix[3][2] 0.073
_struct_site_view.rot_matrix[3][3] -0.172
_struct_site_view.details
; This view highlights the site of ATAT-Netropsin
interaction.
;
```

**\_struct\_site\_view.details** (text)

A description of special aspects of this view of the site. This data item can be used as a figure legend.

Example:

```
; The active site has been oriented with the specificity
pocket on the right and the active site machinery on the left.
; [struct_site_view]
```

**\*\_struct\_site\_view.id** (line)

The value of `_struct_site_view.id` must uniquely identify a record in the `STRUCT_SITE_VIEW` list. Note that this item need not be a number; it can be any unique identifier.

Examples: 'Figure 1', 'unliganded enzyme',

```
'view down enzyme active site'. [struct_site_view]
```

**\_struct\_site\_view.rot\_matrix[1][1]** (float)  
 The [1][1] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_SITE\_GEN category to an orientation useful for visualizing the site. The conventions used in the rotation are described in **\_struct\_site\_view.details**.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_site\_view]

**\_struct\_site\_view.rot\_matrix[1][2]** (float)  
 The [1][2] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_SITE\_GEN category to an orientation useful for visualizing the site. The conventions used in the rotation are described in **\_struct\_site\_view.details**.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_site\_view]

**\_struct\_site\_view.rot\_matrix[1][3]** (float)  
 The [1][3] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_SITE\_GEN category to an orientation useful for visualizing the site. The conventions used in the rotation are described in **\_struct\_site\_view.details**.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_site\_view]

**\_struct\_site\_view.rot\_matrix[2][1]** (float)  
 The [2][1] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_SITE\_GEN category to an orientation useful for visualizing the site. The conventions used in the rotation are described in **\_struct\_site\_view.details**.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_site\_view]

**\_struct\_site\_view.rot\_matrix[2][2]** (float)  
 The [2][2] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_SITE\_GEN category to an orientation useful for visualizing the site. The conventions used in the rotation are described in **\_struct\_site\_view.details**.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_site\_view]

**\_struct\_site\_view.rot\_matrix[2][3]** (float)  
 The [2][3] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_SITE\_GEN category to an orientation useful for visualizing the site. The conventions used in the rotation are described in **\_struct\_site\_view.details**.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_site\_view]

**\_struct\_site\_view.rot\_matrix[3][1]** (float)  
 The [3][1] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_SITE\_GEN category to an orientation useful for visualizing the site. The conventions used in the rotation are described in **\_struct\_site\_view.details**.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_site\_view]

**\_struct\_site\_view.rot\_matrix[3][2]** (float)  
 The [3][2] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_SITE\_GEN category to an orientation useful for visualizing the site. The conventions used in the rotation are described in **\_struct\_site\_view.details**.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_site\_view]

**\_struct\_site\_view.rot\_matrix[3][3]** (float)  
 The [3][3] element of the matrix used to rotate the subset of the Cartesian coordinates in the ATOM\_SITE category identified in the STRUCT\_SITE\_GEN category to an orientation useful for visualizing the site. The conventions used in the rotation are described in **\_struct\_site\_view.details**.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}_{\text{reoriented Cartesian}} = \begin{pmatrix} 11 & 12 & 13 \\ 21 & 22 & 23 \\ 31 & 32 & 33 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{Cartesian}} .$$

[struct\_site\_view]

\* **\_struct\_site\_view.site\_id**

This data item is a pointer to **\_struct\_site.id** in the STRUCT\_SITE category.

SYMMETRY	
Data items in the SYMMETRY category record details about the space-group symmetry.	
Category group(s): <b>inclusive_group</b> <b>symmetry_group</b>	
Category key(s): <b>_symmetry.entry_id</b>	
<i>Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.</i>	
<b>_symmetry.entry_id</b>	'5HVP'
<b>_symmetry.cell_setting</b>	orthorhombic
<b>_symmetry.Int_Tables_number</b>	18
<b>_symmetry.space_group_name_H-M</b>	'P 21 21 2'

## SYMMETRY

### `_symmetry.cell_setting`

`_symmetry_cell_setting` (cif.core.dic 2.0.1)

The cell settings for this space-group symmetry.

The data value must be one of the following:

```
triclinic
monoclinic
orthorhombic
tetragonal
rhombohedral
trigonal
hexagonal
cubic
```

[symmetry]

### \* `_symmetry.entry_id`

This data item is a pointer to `_entry.id` in the ENTRY category.

### `_symmetry.Int_Tables_number`

(int)

`_symmetry_Int_Tables_number` (cif.core.dic 2.0.1)

Space-group number from *International Tables for Crystallography* Vol. A (2002).

[symmetry]

### `_symmetry.space_group_name_H-M`

(line)

`_symmetry_space_group_name_H-M` (cif.core.dic 2.0.1)

Hermann–Mauguin space-group symbol. Note that the Hermann–Mauguin symbol does not necessarily contain complete information about the symmetry and the space-group origin. If used, always supply the *full* symbol from *International Tables for Crystallography* Vol. A (2002) and indicate the origin and the setting if it is not implicit. If there is any doubt that the equivalent positions can be uniquely deduced from this symbol, specify the `_symmetry_equiv.pos_as_xyz` or `_symmetry.space_group_name_Hall` data items as well. Leave spaces between symbols referring to different axes.

Examples: 'P 1 21/m 1', 'P 2/n 2/n 2/n (origin at -1)', 'R -3 2/m'.

[symmetry]

### `_symmetry.space_group_name_Hall`

(line)

`_symmetry_space_group_name_Hall` (cif.core.dic 2.0.1)

Space-group symbol as described by Hall (1981). This symbol gives the space-group setting explicitly. Leave spaces between the separate components of the symbol.

Reference: Hall, S. R. (1981). *Acta Cryst.* **A37**, 517–525; erratum (1981), **A37**, 921.

Examples: '-P 2ac 2n', '-R 3 2"', 'P 61 2 2 (0 0 -1)'.

[symmetry]

## SYMMETRY\_EQUIV

Data items in the SYMMETRY\_EQUIV category list the symmetry-equivalent positions for the space group.

Category group(s): `inclusive_group`

`symmetry_group`

Category key(s): `_symmetry_equiv.id`

*Example 1 – based on PDB entry 5HVP and laboratory records for the structure corresponding to PDB entry 5HVP.*

```
loop_
_symmetry_equiv.id
_symmetry_equiv.pos_as_xyz
1 '+x,+y,+z'
2 '-x,-y,z'
3 '1/2+x,1/2-y,-z'
4 '1/2-x,1/2+y,-z'
```

## 4. DATA DICTIONARIES

## mmcif\_std.dic

### (ucode) \* `_symmetry_equiv.id`

(code)

`_symmetry_equiv_pos_site_id` (cif.core.dic 2.0.1)

The value of `_symmetry_equiv.id` must uniquely identify a record in the SYMMETRY\_EQUIV category. Note that this item need not be a number; it can be any unique identifier.

[symmetry\_equiv]

### `_symmetry_equiv.pos_as_xyz`

(line)

`_symmetry_equiv_pos_as_xyz` (cif.core.dic 2.0.1)

Symmetry-equivalent position in the 'xyz' representation. Except for the space group *P1*, these data will be repeated in a loop. The format of the data item is as per *International Tables for Crystallography* Vol. A (2002). All equivalent positions should be entered, including those for lattice centring and a centre of symmetry, if present.

Example: '-y+x, -y, 1/3+z'.

[symmetry\_equiv]

## VALENCE\_PARAM

Data items in the VALENCE\_PARAM category define the parameters used for calculating bond valences from bond lengths. In addition to the parameters, a pointer is given to the reference (in VALENCE\_REF) from which the bond-valence parameters were taken.

Category key(s): `_valence_param.atom_1`  
`_valence_param.atom_1_valence`  
`_valence_param.atom_2`  
`_valence_param.atom_2_valence`

*Example 1 – a bond-valence parameter list with accompanying references.*

```
loop_
_valence_param.atom_1
_valence_param.atom_1_valence
_valence_param.atom_2
_valence_param.atom_2_valence
_valence_param.Ro
_valence_param.B
_valence_param.ref_id
_valence_param.details
Cu 2 O -2 1.679 0.37 a .
Cu 2 O -2 1.649 0.37 j .
Cu 2 N -3 1.64 0.37 m '2-coordinate N'
Cu 2 N -3 1.76 0.37 m '3-coordinate N'
loop_
_valence_ref.id
_valence_ref.reference
a 'Brown & Altermatt (1985), Acta Cryst. B41, 244-247'
j 'Liu & Thorp (1993), Inorg. Chem. 32, 4102-4205'
m 'See, Krause & Strub (1998), Inorg. Chem. 37, 5369-5375'
```

### \* `_valence_param.atom_1`

(code)

`_valence_param_atom_1` (cif.core.dic 2.3)

The element symbol of the first atom forming the bond whose bond-valence parameters are given in this category.

[valence\_param]

### \* `_valence_param.atom_1_valence`

(int)

`_valence_param_atom_1_valence` (cif.core.dic 2.3)

The valence (formal charge) of the first atom whose bond-valence parameters are given in this category.

[valence\_param]

### \* `_valence_param.atom_2`

(code)

`_valence_param_atom_2` (cif.core.dic 2.3)

The element symbol of the second atom forming the bond whose bond-valence parameters are given in this category.

[valence\_param]

<p><b>*_valence_param.atom_2_valence</b> (int)  _<b>valence_param_atom_2_valence</b> (cif_core.dic 2.3)  The valence (formal charge) of the second atom whose bond-valence parameters are given in this category.</p>	<p><b>_valence_param.Ro</b> (float)  _<b>valence_param_Ro</b> (cif_core.dic 2.3)  The bond-valence parameter <math>R_o</math> used in the expression</p>
[valence_param]	$s = \exp[(R_o - R)/B],$
<p><b>_valence_param.B</b> (float)  _<b>valence_param_B</b> (cif_core.dic 2.3)  The bond-valence parameter <math>B</math> used in the expression</p>	<p>where <math>s</math> is the valence of a bond of length <math>R</math>.</p>
[valence_param]	$s = \exp[(R_o - R)/B],$
<p><b>_valence_param.details</b> (text)  _<b>valence_param_details</b> (cif_core.dic 2.3)  Details of or comments on the bond-valence parameters.</p>	<div style="border: 1px solid black; padding: 5px; text-align: center;"> <p><b>VALENCE_REF</b></p> <p>Data items in the VALENCE_REF category list the references from which the bond-valence parameters have been taken.  Category key(s): <b>_valence_ref.id</b></p> </div>
[valence_param]	
<p><b>_valence_param.id</b> (code)  _<b>valence_param_id</b> (cif_core.dic 2.3)  An identifier for the valence parameters of a bond between the given atoms.</p>	<p><b>*_valence_ref.id</b> (code)  _<b>valence_ref_id</b> (cif_core.dic 2.3)  An identifier for items in this category. Parent of <b>_valence_param.ref_id</b>, which must have the same value.</p>
[valence_param]	[valence_ref]
<p><b>_valence_param.ref_id</b> (code)  _<b>valence_param_ref_id</b> (cif_core.dic 2.3)  An identifier which links to the reference to the source from which the bond-valence parameters are taken. A child of <b>_valence_ref.id</b> which it must match.</p>	<p><b>_valence_ref.reference</b> (text)  _<b>valence_ref_reference</b> (cif_core.dic 2.3)  Literature reference from which the valence parameters identified by <b>_valence_param.id</b> were taken.</p>
[valence_param]	[valence_ref]