CSP CIF dictionary v0.8 (7th CSP Blind Test)

Table 1: CSP CIF fields for use in *method_summary_0-8.cif file*

csp_simulation	Type	Definition	Constraints	Units	Example
_ccdc_csp_simulation_type	char	Indication of how the crystal modelling handles thermal effects.	• static : Simulation does not include thermal corrections of the lattice energy. • dynamic : Simulation includes thermal corrections of the lattice energy.		
csp_conformer_generation	Type	Definition	Constraints	Units	Example
_ccdc_csp_conformer_generation_method	char	Method used for conformer generation. Ideally this should be one of the examples given but other values are permitted.			• Molecular Dynamics • Random Search • Simulated Annealing • Genetic Algorithm • Torsion Scans • Monte Carlo Sampling • Eigenvalues Search • Fragment-based • Rigid Molecule
_ccdc_csp_conformer_generation_software	char	Name of software used for conformer generation.			Conformer Generator (CSD Software)
_ccdc_csp_conformer_generation_software_version	char	Version of software used for conformer generation.			2020.1
_ccdc_csp_conformer_generation_stage	char	Indicates whether this is the final or a	• P : Preliminary conformer generation		

		preliminary method used for conformer generation.	method. There may be more than one. • F : Final conformer generation method. There should only be one.		
csp_conformer_optimisation	Type	Definition	Constraints	Units	• Example
_ccdc_csp_conformer_optimisation_method	char	Method used for conformer optimisation. Ideally this should be one of the examples given but other values are permitted.			• Force Field • Semi-empirical • DFT • Wavefunction • AI
_ccdc_csp_conformer_optimisation_force_field_name	char	Name of force field used for conformer optimisation.			• COMPASSII • CHARM • Williams'98
_ccdc_csp_conformer_optimisation_force_field_description	char	High-level description of the force field used for conformer optimisation.			• Intermolecular only, atomic multiples, exp-6 • Intramolecular and intermolecular, point charges, empirical fitting • AI fitted to DFT structures and energies
_ccdc_csp_conformer_optimisation_semi_empirical_method	char	Semi-empirical method used for conformer optimisation.			• PPP • CNDO/2 • AM1 • PM7

					• Sparkle/AM1
_ccdc_csp_conformer_optimisation_DFT_approximation	char	Type of DFT approximation used for conformer optimisation.			• LDA • GGA • Hybrid • meta-GGA
_ccdc_csp_conformer_optimisation_DFT_functional	char	DFT functional used for conformer optimisation.			• B3LYP • PBE • HS06 • PBE0
_ccdc_csp_conformer_optimisation_DFT_dispersion_correction	char	DFT dispersion correction used for conformer optimisation.			• None • D2 • D3 • TS • MBD
_ccdc_csp_conformer_optimisation_DFT_basis_set	char	DFT basis set used for conformer optimisation.			• atom-centered 6-31G(d,p) • atom-centered cc-pVTZ • numerical DNP • numerical DN
_ccdc_csp_conformer_optimisation_wavefunction_electronic_method	char	Electronic method used by wavefunction for conformer optimisation.			• MP2 • CCSD(T)
_ccdc_csp_conformer_optimisation_wavefunction_basis_set	char	Basis set used by wavefunction for conformer optimisation.			• atom-centered 6-31G(d,p) • atom-centered cc-pVTZ • numerical DNP

					• numerical DN
_ccdc_csp_conformer_optimisation_software	char	Name of software used for conformer optimisation.			• Gaussian16
_ccdc_csp_conformer_optimisation_software_version	char	Version of software used for conformer optimisation.			• Revision B.01
_ccdc_csp_conformer_optimisation_stage	char	Indicates whether this is the final or a preliminary method used for conformer optimisation.	• P : Preliminary conformer optimisation method. There may be more than one. • F : Final conformer optimisation method. There should only be one.		
csp_conformer_clustering	Type	Definition	Constraints	Units	Example
_ccdc_csp_conformer_clustering_method	char	Method used for conformer clustering. Ideally this should be one of the examples given but other values are permitted.			• Energy • Number of conformers • Gyroscopic radius • RMSD • Fingerprint match • CCDC Conformer Generator • Tanimoto similarity • Radial Distribution Function • None
_ccdc_csp_conformer_clustering_cutoff	numb	The conformer clustering cut-off value related to the			• 30 • 200 • 5.0

		clustering method given in _ccdc_csp_conformer_clustering_method.			• 1.0 • 50 • 1.0 • 7.0
_ccdc_csp_conformer_clustering_cutoff_units	char	The units of the conformer clustering cut-off value given in _ccdc_csp_conformer_clustering_cutoff.			• kJ/mol • number of conformers • angstroms
_ccdc_csp_conformer_clustering_software	char	Name of software used for conformer clustering.			• Maestro
_ccdc_csp_conformer_clustering_software_version	char	Version of software used for conformer clustering.			• Schr\"odinger Release 2020-1
_ccdc_csp_conformer_clustering_stage	char	Indicates whether this is the final or a preliminary method used for conformer clustering.	• P : Preliminary conformer clustering method. There may be more than one. • F : Final conformer clustering method. There should only be one.		
csp_structure_generation	Type	Definition	Constraints	Units	Example
_ccdc_csp_structure_generation_method	char	Method used for structure generation. Ideally this should be one of the examples given but other			• Evolutionary algorithm • Random sampling • Simulated annealing • Monte Carlo sampling

		values are permitted.			• Quasi-random sampling
_ccdc_csp_structure_generation_space_group_selection	char	The selection of space groups used for structure generation.	• all : All standard space groups • subset : A subset of standard space groups		
_ccdc_csp_structure_generation_space_group_number_list	numb list	The specific space groups used for structure generation.	• 1:230		
_ccdc_csp_structure_generation_software	char	Name of software used for structure generation.			• Materials Studio • CrystalPredictor
_ccdc_csp_structure_generation_software_version	char	Version of software used for structure generation.			• 2020.1
_ccdc_csp_structure_generation_stage	char	Indicates whether this is the final or a preliminary method used for structure generation.	• P : Preliminary structure generation method. There may be more than one. • F : Final structure generation method. There should only be one.		
csp_structure_optimisation	Type	Definition	Constraints	Units	Example
_ccdc_csp_structure_optimisation_method	char	Method used for structure optimisation. Ideally this should be one of the examples given but other			• Force Field • Semi-empirical • DFT • Wavefunction • AI

		values are permitted.			
_ccdc_csp_structure_optimisation_force_field_name	char	Name of force field used for structure optimisation.			• COMPASSII • CHARM • Williams'98
_ccdc_csp_structure_optimisation_force_field_description	char	High-level description of the force field used for structure optimisation.			• Intermolecular only, atomic multiples, exp-6 • Intramolecular and intermolecular, point charges, empirical fitting • AI fitted to DFT structures and energies
_ccdc_csp_structure_optimisation_semi_empirical_method	char	Semi-empirical method used for structure optimisation.			• PPP • CNDO/2 • AM1 • PM7 • Sparkle/AM1
_ccdc_csp_structure_optimisation_DFT_approximation	char	Type of DFT approximation used for structure optimisation.			• LDA • GGA • Hybrid • meta-GGA
_ccdc_csp_structure_optimisation_DFT_functional	char	DFT functional used for structure optimisation.			• B3LYP • PBE • HS06 • PBE0
_ccdc_csp_structure_optimisation_DFT_dispersion_correction	char	DFT dispersion correction used for			• None • D2

		structure optimisation.			• D3 • TS • MBD
_ccdc_csp_structure_optimisation_DFT_basis_set	char	DFT basis set used for structure optimisation.			• atom-centered 6-31G(d,p) • atom-centered cc-pVTZ • numerical DNP • numerical DN
_ccdc_csp_structure_optimisation_wavefunction_electronic_method	char	Electronic method used by wavefunction for structure optimisation.			• MP2 • CCSD(T)
_ccdc_csp_structure_optimisation_wavefunction_basis_set	char	Basis set used by wavefunction for structure optimisation.			• atom-centered 6-31G(d,p) • atom-centered cc-pVTZ • numerical DNP • numerical DN
_ccdc_csp_structure_optimisation_software	char	Name of software used for structure optimisation.			• Gaussian16
_ccdc_csp_structure_optimisation_software_version	char	Version of software used for structure optimisation.			• Revision B.01
_ccdc_csp_structure_optimisation_stage	char	Indicates whether this is the final or a preliminary method used for structure optimisation.	• P : Preliminary structure optimisation method. There may be more than one.		

			• F : Final structure optimisation method. There should only be one.		
csp_structure_clustering	Type	Definition	Constraints	Units	• Example
_ccdc_csp_structure_clustering_method	char	Method used for structure clustering. Ideally this should be one of the examples given but other values are permitted.			• Energy • Number of structures • Molecule cluster overlay • pXRD similarity • Fingerprint match • RDF (Radial Distribution Function) • None
_ccdc_csp_structure_clustering_cutoff	numb	The structure clustering cut-off value related to the clustering method given in _ccdc_csp_structure _clustering_method.			• 50 • 100 • 25 • 0.8 • 200 • 10.0
_ccdc_csp_structure_clustering_cutoff_ units	char	The units of the structure clustering cut-off value given in _ccdc_csp_structure _clustering_cutoff.			• kJ/mol • number of structures • molecules matched • angstroms
_ccdc_csp_structure_clustering_ software	char	Name of software used for structure clustering.			• COMPACK

_ccdc_csp_structure_clustering_ software_version	char	Version of software used for structure clustering.			• 2020.1
_ccdc_csp_structure_clustering_stage	char	Indicates whether this is the final or a preliminary method used for structure clustering.	• P : Preliminary structure clustering method. There may be more than one. • F : Final structure clustering method. There should only be one.		
csp_free_energy	Type	Definition	Constraints	Units	• Example
_ccdc_csp_free_energy_correction_ method	char	Method used for free energy correction. Ideally this should be one of the examples given but other values are permitted.			• Harmonic rigid-molecule • Harmonic all-atoms • Anharmonic rigid-molecule • Anhamonic all-atoms • Quasi-harmonic rigid-molecule • Quasi-harmonic all-atoms • Thermal averaging • Zero-point energy rigid-molecule • Zero-point energy all-atoms
_ccdc_csp_free_energy_correction_ software	char	Name of software used for free energy correction.			• DMACRYS • DPOLY

_ccdc_csp_free_energy_correction_ software_version	char	Version of software used for free energy correction.			
_ccdc_csp_free_energy_correction_ stage	char	Indicates whether this is the final or a preliminary method used for free energy correction.	• P : Preliminary free energy correction method. There may be more than one. • F : Final free energy correction method. There should only be one.		
csp_reference	Type	Definition	Constraints	Units	• Example
_ccdc_csp_reference_category	char	The category of method described by a publication reference. Ideally this should be one of the method categories listed but other values are permitted.			• Conformer generation • Conformer optimisation • Conformer clustering • Structure generation • Structure optimisation • Structure clustering • Free energy correction • General
_ccdc_csp_reference_citation	char	Citation for an article or other publication that describes one or more of the methods used for this prediction.			• J. C. Cole, C. R. Groom, M. G. Read, I. Giangreco, P. McCabe, A. M. Reilly, G. P. Shields (2016). Generation of crystal structures using known crystal structures as

					analogues. Acta Crystallogr Sect B Struct Sci Cryst Eng Mater. 72(4):530-541.
_ccdc_csp_reference_identifier	char	A persistent identifier (for example a DOI) indicating the location of an article or other publication that decribes one or more of the methods used for this prediction.			• https://doi.org/10.1107/S2052520616006533
csp_audit	Type	Definition	Constraints	Units	Example
_ccdc_csp_audit_blind_test_number	numb	Number of the Blind Test for which this result was submitted.	7		• 7
_ccdc_csp_audit_blind_test_structure_label	char	Label of the Blind Test structure for which this result was submitted.	XXVII XXVIII XXIX XXX XXXI XXXII XXXIII		
_ccdc_csp_audit_blind_test_landscape_label	char	Label of the landscape to which this data block relates. This combined with other CSP audit fields should allow the identification of all data associated with a particular landscape.			

CSP CIF dictionary v0.8 (7th CSP Blind Test)

Contributors (loop)	Type	Definition	Constraints	Units	Example
_publ_author_name _publ_author_address _publ_author_id_orcid		https://www.iucr.org/__data/iucr/cifdic_html/1/cif_core.dic/Cpubl_author.html			
Main Contact	Type	Definition	Constraints	Units	Example
_audit_contact_author_name _audit_contact_author_address _audit_contact_author_email		https://www.iucr.org/__data/iucr/cifdic_html/1/cif_core.dic/Caudit_contact_author.html			

Table 2: CSP CIF fields included in *structure_specific_data_0-8.csv file (values to be included under corresponding columns using excel or similar program)*

csp_density_calc	Type	Definition	Constraints	Units	Example	Obligation
_ccdc_csp_density_calc	numb	Density value calculated from the predicted crystal cell and contents.	0.0:	megagrams per cubic metre (Mgm^-3^)	• 1.532	
csp_simulation	Type	Definition	Constraints	Units	Example	Obligation
_ccdc_csp_simulation_temperature	numb	The temperature at which the structure is calculated.	0.0:	kelvin (K)	• 298.15	Required
_ccdc_csp_simulation_pressure	numb	The pressure at which the structure is calculated.	0.0:	kilopascals (kPa)		
csp_conformer_clustering	Type	Definition	Constraints	Units	Example	Obligation

_ccdc_csp_conformer_clustering_identifier	char	External identifier or label of a conformer used in a CSP calculation.			• conf_1	
csp_classification	Type	Definition	Constraints	Units	Example	Obligation
_ccdc_csp_classification_energy_absolute	numb	The total (absolute) energy of the crystal structure, with respect to gas-phase atoms.		electron volt (eV)	• -16583.5564687	Desirable
_ccdc_csp_classification_energy_lattice_relative	numb	The relative lattice energy of the structure with respect to the global minimum on the lattice or absolute energy landscape.		kilojoule per mole (kJ/mol)	• 7.3	Required
_ccdc_csp_classification_energy_lattice_absolute	numb	Lattice energy, with respect to isolated gas-phase molecules in their lowest energy conformation.		kilojoule per mole (kJ/mol)	• -250.345	
_ccdc_csp_classification_ranking	numb	The rank of the structure when ordered by e.g. increasing relative energy or other chosen criteria. 1 is the top ranked.			• 3	Required

_ccdc_csp_classification_energy_error_absolute	numb	Estimate of the signed absolute lattice energy error evaluated as (predicted - experiment).		kilojoule per mole (kJ/mol)	• 12.5	
_ccdc_csp_classification_energy_error_relative	numb	Estimate of the signed relative lattice energy error defined as ((predicted - experimental)/experimental) and given as a simple ratio (not a percentage).	0.0:1.0		• 0.049	
csp_free_energy	Type	Definition	Constraints	Units	Example	Obligation
_ccdc_csp_free_energy_correction_absolute	numb	The free energy correction of the total (absolute) energy at a given temperature.		kilojoule per mole (kJ/mol)	• 4.2	Desirable
_ccdc_csp_free_energy_correction_lattice	numb	The free energy correction of the lattice energy at a given temperature.		kilojoule per mole (kJ/mol)		Desirable
_ccdc_csp_free_energy_relative	numb	The relative free energy, at a given temperature, of the structure with respect to the corresponding global minimum on the same free energy landscape.		kilojoule per mole (kJ/mol)	• 7.3	**Required**

_ccdc_csp_free_energy_ranking	numb	The rank of the structure at a given temperature when ordered by relative free energy or free energy correction.			• 2	Required

Table 3: CIF fields required within individual CIF files submitted by participants

Chemical Formula	Type	Definition	Constraints	Units	Example	Obligation
_chemical_formula_sum		https://www.iucr.org/__data/iucr/cifdic_html/1/cif_core.dic/Cchemical_formula.html				Required
Space Group	Type	Definition	Constraints	Units	Example	Obligation
_space_group_crystal_system _space_group_name_H-M_alt _space_group_IT_number		https://www.iucr.org/__data/iucr/cifdic_html/1/cif_core.dic/Cspace_group.html				Required
Symmetry Operators (loop)	Type	Definition	Constraints	Units	Example	Obligation
_symmetry_equiv_pos_site_id _symmetry_equiv_pos_as_xyz		https://www.iucr.org/__data/iucr/cifdic_html/1/cif_core.dic/Csymmetry_equiv.html				Required
Cell Parameters	Type	Definition	Constraints	Units	Example	Obligation
_cell_length_a _cell_length_b		https://www.iucr.org/__data/iucr/cifdic				Required

_cell_length_c _cell_angle_alpha _cell_angle_beta _cell_angle_gamma _cell_volume _cell_formula_units_Z		_html/1/cif_core.dic/Ccell.html				
Atomic Coordinates (loop)	Type	Definition	Constraints	Units	Example	Obligation
_atom_site_label _atom_site_type_symbol _atom_site_fract_x _atom_site_fract_y _atom_site_fract_z		https://www.iucr.org/__data/iucr/cifdic_html/1/cif_core.dic/Catom_site.html				Required
Creation Date	Type	Definition	Constraints	Units	Example	Obligation
_audit_creation_date		https://www.iucr.org/__data/iucr/cifdic_html/1/cif_core.dic/Iaudit_creation_date.html				Desirable