ID	PCCD Ligand Instance ID$_{2008}$	PCCD Ligand Instance ID$_{2010}$	PDB ID	Ligand ID	Chain ID	Residue	mmCIF Asym ID	mmCIF Sequence No.	Disorder Flag	Model No.	$B_{average}$	B_{max}	B_{min}	B_{stddev}	$E_{solvent}$	$E_{solvent,H}$	$E_{solvent,len}$	$E_{solvent,ang}$	$E_{solvent,tor}$	$\Delta E_{solvent,ang\text{-}tor}$
783	149_1jz5_1__2001__E__	149_1jz5_1_A_2001__E__	1jz5	149	A	2001	E			1	21.1667	28.07	16.45	3.69347	5.90196	3.70568	-2.8907	-4.85041	-24.2988	19.44839
1195	19A_2ojg_1__301__C__	19A_2ojg_1_A_360__C__	2ojg	19A	A	360	C			1	47.0833	51.9	41.61	2.91662	14.8927	11.7263	6.1109	5.02384	-13.29	18.31384
1333	1CA_1y9r_1__2001__C__	1CA_1y9r_1_A_2001__C__	1y9r	1CA	A	2001	C			1	22.2571	27.3	18.52	2.00606	16.0172	11.6735	-28.7588	-39.4183	-82.985	43.5667
2447	1PU_1gii_1__501__B__	1PU_1gii_1_A_501__B__	1gii	1PU	A	501	B			1	34.4461	37.82	28.84	2.18631	11.9971	9.95577	3.00234	1.75046	-15.4301	17.18056
4152	2OS_1t3i_1__602__C__	2OS_1t3i_1_A_602__C__	1t3i	2OS	A	602	C			1	26.8081	38.79	21.29	4.27433						
6153	458_2b1v_1__201__E__	458_2b1v_1_A_201__E__	2b1v	458	A	201	E			1	28.9958	48.4	19.12	7.62622	10.9759	8.24878	0.745648	-2.32538	-22.2457	19.92032
6166	468_2h7p_1__501__C__	468_2h7p_1_A_501__C__	2h7p	468	A	501	C			1	32.1822	35.99	30.03	1.66934	15.6593	12.2005	-104.994	-109.603	-188.532	78.929
6320	4DE_1y2d_1__602__H__	4DE_1y2d_1_B_602__H__	1y2d	4DE	B	602	H			1	46.996	60.67	30.72	9.34249	19.822	16.6945	12.9483	-1.12925	-15.0198	13.89055
6941	566_2h7i_1__401__C__	566_2h7i_1_A_401__C__	2h7i	566	A	401	C			1	31.4476	33.11	29.46	1.01567	12.4582	9.50451	6.21259	5.86558	-34.5379	40.40348
8691	665_2h7l_1__401__C__	665_2h7l_1_A_401__C__	2h7l	665	A	401	C			1	26.6345	32.17	24.11	1.89509	13.0113	10.0319	-6.22936	-18.4005	-124.079	105.6785
8808	6GP_2f3q_1__998__E__	6GP_2f3q_1_A_998__C__	2f3q	6GP	A	998	C			1	25.1439	37.06	17.96	6.39166	12.6135	9.98169	5.64247	-1.91777	-17.2378	15.32003
9467	7GP_2f3s_1__998__E__	7GP_2f3s_1_A_998__C__	2f3s	7GP	A	998	C			1	27.4695	39.23	20.5	6.34385						
10878	A2G_2z48_1_A_1001__C__	A2G_2z48_1_A_1001__C__	2z48	A2G	A	1001	C			1	28.1533	32.66	22.16	2.61353	9.83342	7.1665	6.11103	5.22562	-4.28303	9.50865
11002	A2G_1ax0_1__401__E__	A2G_1ax0_1_A_401__C__	1ax0	A2G	A	401	C			1	22.784	27.97	18.34	2.79787	9.80941	7.13057	6.57585	5.70926	-1.8949	7.60416
11017	A2G_2ccv_1_A_1101__G__	A2G_2ccv_1_A_1101__C__	2ccv	A2G	A	1101	C			1	10.58	12.87	9.54	0.964393	10.0046	7.18142	3.2871	0.00710529	-9.58187	9.58897529
11034	A2G_2jhi_1_F_1299__B__	A2G_2jhi_1_F_1299__C__	2jhi	A2G	F	1299	C			1	36.7347	40.83	31.84	2.57151	10.1279	7.5432	6.61386	5.59416	-3.50285	9.09701
33545	AI1_1fd7_1_D_104__K__	AI1_1fd7_1_D_104__K__	1fd7	AI1	D	104	K			1	33.9521	43.38	25.46	6.69438						
35593	AMG_2z49_1_A_1400__C__	AMG_2z49_1_A_1400__C__	2z49	AMG	A	1400	C			1	25.3015	30.87	21.5	2.24844	8.87816	6.4679	5.44568	4.63369	-3.51201	8.1457
38193	AOS_2i57_1__3001__E__	AOS_2i57_1_A_3001__E__	2i57	AOS	A	3001	E			1	31.265	39.34	18.84	5.60378						
43086	B1V_1tou_1__501__B__	B1V_1tou_1_A_501__B__	1tou	B1V	A	501	B			1	31.18	36.21	29.19	2.08142	11.6749	8.92262	1.02855	-2.44597	-24.7018	22.25583
45987	BDF_1ovp_1__115__E__	BDF_1ovp_1_A_115__B__	1ovp	BDF	A	115	B			1	8.91167	13.52	6.61	1.8504	6.7525	4.50978	-9.00489	-10.076	-34.965	24.889
45988	BDF_2fa1_1__1__C__	BDF_2fa1_1_B_1__C__	2fa1	BDF	B	1	C			1	18.1842	20.67	16.32	1.44639	6.2523	3.68139	1.1952	-2.32137	-26.0902	23.76883
47649	BGC_2dqa_1_B_2001__E__	BGC_2dqa_1_B_2001__E__	2dqa	BGC	B	2001	E			1	39.415	40.95	35.61	1.38952	5.70722	3.64648	3.27374	3.17585	-2.50939	5.68524
47652	BGC_2e9l_1_A_1401__B__	BGC_2e9l_1_A_1401__B__	2e9l	BGC	A	1401	B			1	21.0667	24.53	16.16	2.47231	5.84365	3.712	3.26333	3.00041	-10.0857	13.08611
48095	BGC_2b3b_1__2607__I__	BGC_2b3b_1_C_2607__I__	2b3b	BGC	C	2607	I			1	14.5658	16.39	12.4	1.24809	5.87806	4.00243	3.03795	2.5874	-3.15996	5.74736
48123	BGC_2cdb_1_B_1368__F__	BGC_2cdb_1_B_1368__K__	2cdb	BGC	B	1368	K			1	13.72	18.03	10.66	2.28245						
48149	BGC_2h3h_1__1500__C__	BGC_2h3h_1_A_1500__C__	2h3h	BGC	A	1500	C			1	11.7683	14.5	10.09	1.13464	5.88498	3.90582	3.14089	2.77756	-5.99816	8.77572
48158	BGC_2ipl_1__501__B__	BGC_2ipl_1_A_501__B__	2ipl	BGC	A	501	B			1	7.82583	9.09	6.98	0.549749	5.64706	3.6679	3.31022	2.71597	-2.92785	5.64382
48159	BGC_2ipm_1__501__B__	BGC_2ipm_1_A_501__B__	2ipm	BGC	A	501	B			1	4.4725	5.39	3.86	0.423519	5.58953	3.58087	3.14099	2.72181	-2.53734	5.25915
48160	BGC_2ipn_1__701__B__	BGC_2ipn_1_A_701__B__	2ipn	BGC	A	701	B			1	5.78083	6.39	5.27	0.400114	5.6073	3.61998	3.32003	2.73394	-2.55722	5.29116
48193	BGC_2pok_1__2__D__	BGC_2pok_1_A_459__D__	2pok	BGC	A	459	D			1	35.22	37.65	31.39	1.78871	5.87343	3.93694	3.00759	2.28972	-3.58	5.86972
48196	BGC_2pwf_1_E_9998__E__	BGC_2pwf_1_A_9998__E__	2pwf	BGC	A	9998	E			1	11.2317	17.17	8.75	2.46952	5.75869	3.73498	3.24803	3.17712	-2.83505	6.01217
48577	BHG_1lzi_1__452__B__	BHG_1lzi_1_A_452__B__	1lzi	BHG	A	452	B			1	16.4594	27.01	12.82	4.42448						
48578	BHG_1lzj_1__452__B__	BHG_1lzj_1_A_452__B__	1lzj	BHG	A	452	B			1	19.9628	28.98	16.08	4.16579	13.3264	9.98364	5.04451	-1.63658	-15.1713	13.53472
48633	BHL_1xvb_1__1202__M__	BHL_1xvb_1_B_1202__M__	1xvb	BHL	B	1202	M			1	44.42	44.42	44.42	7.11E-15	6.37756	4.31117	-2.43455	-21.1726	-35.0155	13.8429
48706	BI5_1zyj_1__201__C__	BI5_1zyj_1_A_362__C__	1zyj	BI5	A	362	C			1	39.7826	47.75	34.38	4.86439	14.6971	12.3157	-29.5435	-31.0257	-41.4279	10.4022
50128	BMA_2zbl_1_A_501__G__	BMA_2zbl_1_A_501__G__	2zbl	BMA	A	501	G			1	9.38833	10.48	7.82	0.721663	5.70832	3.73292	3.32128	2.88327	-3.23243	6.1157
54699	BOG_2qcu_1_A_700__C__	BOG_2qcu_1_A_700__C__	2qcu	BOG	A	700	C			1	110.705	116.1	99.37	5.3242						
54732	BOG_3b6h_1_A_701__C__	BOG_3b6h_1_A_701__C__	3b6h	BOG	A	701	C			1	46.4	50.94	37.45	5.05728	14.9753	11.0559	9.48649	6.5623	-14.2585	20.8208
55291	BP3_1lgt_1__301__GA__	BP3_1lgt_1_A_301__E__	1lgt	BP3	A	301	E			1	26.4053	29.21	23.64	1.79397	8.94007	7.42024	1.97345	-0.941954	-3.15706	2.215106
55300	BP6_1lkd_1__301__GA__	BP6_1lkd_1_A_301__E__	1lkd	BP6	A	301	E			1	37.2713	38.3	36.05	0.694297	9.79659	8.38971	2.29032	-0.46821	-3.36479	2.89658
55728	BPY_2ei3_1_A_351__E__	BPY_2ei3_1_A_351__E__	2ei3	BPY	A	351	E			1	45.1521	46.91	44.41	0.627765	7.45638	5.84306	5.8236	3.30729	1.16309	2.1442
55730	BPY_1eir_1_A_301__Q__	BPY_1eir_1_A_301__C__	1eir	BPY	A	301	C			1	26.4079	31.96	20.44	4.06077	8.70275	6.98338	-15.7199	-18.0455	-20.2844	2.2389
57837	BRJ_1xvg_1__3800__O__	BRJ_1xvg_1_B_3800__O__	1xvg	BRJ	B	3800	O			1	43.78	43.78	43.78	0	3.56295	2.74907	2.16737	1.48338	-2.09028	3.57366
57849	BRJ_2f4g_1__300__E__	BRJ_2f4g_1_A_300__E__	2f4g	BRJ	A	300	E			1	31.915	36.82	23.4	5.1964	3.54576	2.5361	1.4913	-13.1485	-22.841	9.6925

Column ID	Column name	Originating from	Explanation
	Values identifying ligand instance:		
A	ID	our own assignment	sequential assignment of numerical ID
B	PCCD Ligand Instance ID_{2008}	Ligand Expo (see Note)	Ligand Instance ID from 2008 download (Note: In our internal nomenclature, we inverted the sequence of substrings PDB ID and Ligand ID)
C	PCCD Ligand Instance ID_{2010}	Ligand Expo (see Note)	Ligand Instance ID from 2010 download (Note: In our internal nomenclature, we inverted the sequence of substrings PDB ID and Ligand ID)
D	PDB ID	Ligand Expo Instance ID	PDB ID
E	Ligand ID	Ligand Expo Instance ID	Chemical Component Identifier (3-letter-code), also known as Ligand ID or HET ID
F	Chain ID	Ligand Expo Instance ID	Chain ID
G	Residue	Ligand Expo Instance ID	Residue ID number of the ligand instance
H	mmCIF Asym ID	Ligand Expo Instance ID	mmCIF Asym ID
I	mmCIF Sequence No.	Ligand Expo Instance ID	mmCIF sequence number
J	Disorder Flag	Ligand Expo Instance ID	Disorder Flag (for structures modeled with disorder)
K	Model No.	Ligand Expo Instance ID	Model number
	B value statistics:		
L	$B_{average}$	Ligand Expo PDBML file	Average B value of ligand instance atoms
M	B_{max}	Ligand Expo PDBML file	Largest B value among ligand instance atoms
N	B_{min}	Ligand Expo PDBML file	Smallest B value among ligand instance atoms
O	B_{stddev}	Ligand Expo PDBML file	Standard deviation of B values of ligand instance atoms
	Solvent model QC energies:		*Values in kcal/mol, relative to vacuum X-ray conformational energy*
P	$E_{solvent}$	G09 solvent model runs	Conformational energy of ligand with unmodified heavy-atom crystal coordinates and hydrogens added by CACTVS, calculated in IEFPCM solvent model
Q	$E_{solvent,H}$	G09 solvent model runs	Solvent model conformational energy of ligand after optimization of hydrogens (all other internal coordinates kept fixed)
R	$E_{solvent,len}$	G09 solvent model runs	Solvent model conformational energy of ligand after optimization of bond lenghts between heavy atoms (bond angles and torsions kept fixed)
S	$E_{solvent,ang}$	G09 solvent model runs	Solvent model conformational energy of ligand after optimization of bond angles (torsions kept fixed)
T	$E_{solvent,tor}$	G09 solvent model runs	Solvent model conformational energy of ligand after optimization of torsions (full unconstrained optimization)
U	$\Delta E_{solvent,ang-tor}$	G09 solvent model runs	Energy difference between angle-optimized and torsion-optimized solvent model conformations ([S]-[T]), **"solvent conformational energy"**
	Vacuum QC energies:		*Values in kcal/mol, relative to vacuum X-ray conformational energy*
V	$E_{vacuum,H}$	G09 vacuum runs	Vacuum conformational energy of ligand after optimization of hydrogens (all other internal coordinates kept fixed)
W	$E_{vacuum,len}$	G09 vacuum runs	Vacuum conformational energy of ligand after optimization of bond lenghts between heavy atoms (bond angles and torsions kept fixed)
X	$E_{vacuum,ang}$	G09 vacuum runs	Vacuum conformational energy of ligand after optimization of bond angles (torsions kept fixed)
Y	$E_{vacuum,tor}$	G09 vacuum runs	Vacuum conformational energy of ligand after optimization of torsions (full unconstrained optimization)
Z	$\Delta E_{vacuum,H-len}$	G09 vacuum runs	Energy difference between hydrogen-optimized and bond-length-optimized vacuum conformations ([V]-[W])
AA	$\Delta E_{vacuum,len-ang}$	G09 vacuum runs	Energy difference between bond-length-optimized and angle-optimized vacuum conformations ([W]-[X])
AB	$\Delta E_{vacuum,ang-tor}$	G09 vacuum runs	Energy difference between angle-optimized and torsion-optimized vacuum conformations ([X]-[Y]), **"vacuum conformational energy"**
AC	$\Delta E_{vacuum,len-tor}$	G09 vacuum runs	Energy difference between bond-length-optimized and torsion-optimized vacuum conformations ([W]-[Y])
AD	E_{GEM}	MacroModel & G09 global E min. runs	Vacuum conf. E of lowest of (up to) three QC-reoptimized, OPLS-2001-calculated, lowest-energy ligand conformations ("global" energy minimum)
AE	$\Delta E_{vacuum,len-GEM}$	G09 vacuum runs	Energy difference between bond-length-optimized and "global" energy minimum vacuum conformations ([W]-[AD])
AF	$\Delta E_{vacuum,ang-GEM}$	G09 vacuum runs	Energy difference between angle-optimized and "global" energy minimum vacuum conformations ([X]-[AD])
AG	$\Delta E_{vacuum,tor-GEM}$	G09 vacuum runs	Energy difference between torsion-optimized and "global" energy minimum vacuum conformations ([Y]-[AD])
AH	E_{OE_GEM}	OpenEye Omega runs	Molecular mechanics MMFF94s energy difference between crystal conformation and lowest-energy conformation from Omega conformational search
	Comparison solvent model, vacuum QC energies:		
AI	$\Delta\Delta E_{ang-tor}$	G09 solvent and vacuum runs	Difference between solvent and vacuum conformational energies ([U]-[AB])
	Residue-specific crystallographic "quality" parameters:		
AJ	RSR_{2008}	Manual extraction from EDS late 2008	Real -space R-factor of ligand instance
AK	RSR_{2010}	Parsed from 2010 download from EDS	Real -space R-factor of ligand instance
AL	$RSCC_{2008}$	Manual extraction from EDS late 2008	Real-space correlation coefficient of ligand instance
AM	$RSCC_{2010}$	Parsed from 2010 download from EDS	Real-space correlation coefficient of ligand instance
AN	$OWAB_{2008}$	Manual extraction from EDS late 2008	Occupancy-weighted average B-factor of ligand instance atoms
AO	$OWAB_{2010}$	Parsed from 2010 download from EDS	Occupancy-weighted average B-factor of ligand instance atoms
AP	$OWAB_{calculated}$	Calculated from PDBML file	Occupancy-weighted average B-factor of ligand instance atoms
	Results of torsional sampling around crystal conformationat MMFF94s level:		
AQ	$n_{torsion}$	Maestro runs using MMFF94s	Number of sampling iterations achieved for sampling around ligand crystal conformation using gaussian distribution of torsional perturbations
AR	$E_{torsion,min}$	Maestro runs using MMFF94s	Minimum of MMFF94s energies of torsion-sampled conformations around ligand crystal conformation, using gaussian distribution of torsional perturbations
AS	$E_{torsion,max}$	Maestro runs using MMFF94s	Maximum of MMFF94s energies of torsion-sampled conformations around ligand crystal conformation, using gaussian distribution of torsional perturbations
AT	$E_{torsion,avg}$	Maestro runs using MMFF94s	Arithmetic mean of MMFF94s energies of torsion-sampled conformations around ligand crystal conformation, using gaussian distribution of torsional perturbations
AU	$E_{torsion,median}$	Maestro runs using MMFF94s	Median of MMFF94s energies of torsion-sampled conformations around ligand crystal conformation, using gaussian distribution of torsional perturbations
AV	$E_{torsion,SP}$	Maestro runs using MMFF94s	MMFF94s energy of ligand instance crystal conformation (after optimization of added hydrogens)
AW	$\Delta E_{torsion,avg-SP}$	Maestro runs using MMFF94s	MMFF94s energy difference between arithmetic mean of sampled conformations and ligand crystal conformation ([AT]-[AV])
AX	$\Delta E_{torsion,max-SP}$	Maestro runs using MMFF94s	MMFF94s energy difference between maximum E among sampled conformations and ligand crystal conformation ([AS]-[AV])
AY	$\Delta E_{torsion,SP-min}$	Maestro runs using MMFF94s	MMFF94s energy difference between ligand crystal conformation and minimum E among sampled conformations ([AV]-[AR])
AZ	$\Delta E_{torsion,median-SP}$	Maestro runs using MMFF94s	MMFF94s energy difference between median of sampled conformations and ligand crystal conformation ([AU]-[AV])
	Ligand molecule properties:		*All properties calculated by CACTVS*
BA	n_{rot}	PCCD ligand (prototype) structure	Number of freely rotatable bonds
BB	n_{rot_eff}	PCCD ligand (prototype) structure	Number of freely rotatable bonds + partially accounted macrocyclic single bonds
BC	$n_{tautomers}$	PCCD ligand (prototype) structure	Number of possible tautomers
BD	$n_{hdonors}$	PCCD ligand (prototype) structure	Number of hydrogen bond donors
BE	$n_{hacceptors}$	PCCD ligand (prototype) structure	Number of hydrogen bond acceptors
BF	$n_{ringsystems}$	PCCD ligand (prototype) structure	Number of ring systems
BG	n_{rings}	PCCD ligand (prototype) structure	Number of rings (smallest number of smallest rings)
BH	n_{heavy_atoms}	PCCD ligand (prototype) structure	Number of heavy (non-hydrogen) atoms
BI	$n_{heteroatoms}$	PCCD ligand (prototype) structure	Number of hetero atomes
BJ	$n_{hydrogens}$	PCCD ligand (prototype) structure	Number of hydrogen atoms
BK	aromatic	PCCD ligand (prototype) structure	Boolean: Is molecule aromatic?
BL	macrocyclic	PCCD ligand (prototype) structure	Boolean: Is molecule a macrocycle?
BM	MW	PCCD ligand (prototype) structure	Molecular weight
BN	xlogp2	PCCD ligand (prototype) structure	Octanol/water partition coefficient, calculated by XLOGP v. 2.0 algorithm
	Crystallographic properties of the entire PDB structure (protein & ligands):		*All values taken as is from PDBML files; no cross-check vs. original publication*
BO	Resolution	PDBML file	Crystallographic resolution
BP	R_{free}	PDBML file	R_{free}
BQ	R	PDBML file	R value
BR	R_{free}-R	PDBML file	R_{free}- R
BS	Solvent	PDBML file	Solvent content of crystal
BT	T	PDBML file	Temperature during crystal growth / co-crystallization
BU	$T_{ambient}$	PDBML file	Temperature during data collection
BV	B_{mean}	PDBML file	Average B value (of entire PDB structure)
BW	pH	PDBML file	pH value
BX	Year	PDBML file	Year of publication
BY	EC code	PDBML file	EC code of protein
BZ	Keywords	PDBML file	Keywords associated with PDB structure
CA	Macromolecule	PDBML file	Protein name
CB	Macromolecule Type	PDBML file	Protein type
CC	Method	PDBML file	Experimental method used (here for all: X-ray diffraction)
CD	Organism	PDBML file	Organism that was source of protein (if given)
CE	DPI	PDBML file	Diffraction-component precision index (or: diffraction-data precision indicator) of the experimental data
CF	Space Group	PDBML file	Space group of crystal
CG	Diffraction Source	PDBML file	Source of X-ray beam
CH	Refinement Program	PDBML file	Software used for structure refinement
	Various ligand molecule properties/identifiers:		
CI	Sugar	manual assignment	Boolean: Is ligand molecule a sugar? (Heuristically determined from ligand molecule name [next column].)
CJ	Name	Ligand Expo mmCIF file	Ligand molecule name
CK	OpenEye Name	Ligand Expo mmCIF file	OpenEye ligand molecule name (calculated IUPAC name)
CL	ACD/Labs Name	Ligand Expo mmCIF file	ACD Labs ligand molecule name (calculated IUPAC name)
CM	SMILES	PCCD ligand (prototype) structure	SMILES string of ligand molecule (calculated by CACTVS)
	HQ subset membership:		
CN	VHQ-R subset		member of the “Very High-Quality set with Results” (VHQ-R) subset
CO	UHQ-R subset		member of the “Ultra High-Quality subset with Results” (UHQ-R) subset
	Abbreviations used:		
	EDS	Electron Density Server	Uppsala Electron Density Server, http://eds.bmc.uu.se/eds/
	PCCD	PDB Chemical Component Dictionary	http://deposit.rcsb.org/cc_dict_tut.html
	QC	quantum-chemical (or quantum chemistry)	
	Notes:		
	Greenish entries are used in manuscript figures		