Full wwPDB X-ray Structure Validation Report ⓘ

Feb 16, 2017 – 07:23 am GMT

PDB ID	:	4QL1
Title	:	Crystal structure of human WDR5 in complex with compound OICR-9429
Authors	:	Dong, A.; Dombrovski, L.; Walker, J.R.; Getlik, M.; Kuznetsova, E.; Smil, D.; Barsyte, D.; Li, F.; Poda, G.; Senisterra, G.; Marcellus, R.; Al-Awar, R.; Bountra, C.; Arrowsmith, C.H.; Edwards, A.M.; Brown, P.J.; Schapira, M.; Vedadi, M.; Wu, H.; Structural Genomics Consortium (SGC)
Deposited on	:	2014-06-10
Resolution	:	1.50 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Mogul	:	1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix)	:	1.9-1692
EDS	:	FAILED
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28986

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.50 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

Metric	Percentile Ranks	Value
Clashscore		2
Ramachandran outliers		0
Sidechain outliers		0.4%

Worse   *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
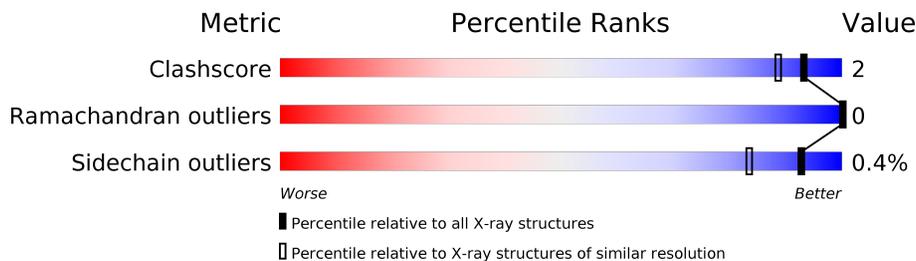
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	112137	2503 (1.50-1.50)
Ramachandran outliers	110173	2445 (1.50-1.50)
Sidechain outliers	110143	2443 (1.50-1.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Note EDS failed to run properly.

Mol	Chain	Length	Quality of chain
1	A	312	95%
1	B	312	92%

2 Entry composition ⓘ

There are 6 unique types of molecules in this entry. The entry contains 5115 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called WD repeat-containing protein 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	304	Total	C	N	O	S	0	6	0
			2346	1501	387	448	10			
1	B	302	Total	C	N	O	S	0	9	0
			2353	1507	385	450	11			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	23	GLY	-	EXPRESSION TAG	UNP P61964
B	23	GLY	-	EXPRESSION TAG	UNP P61964

- Molecule 2 is N-(4-(4-METHYLPIPERAZIN-1-YL)-3'-(MORPHOLINOMETHYL)-[1,1'-BIPHENYL]-3-YL)-6-OXO-4-(TRIFLUOROMETHYL)-1,6-DIHYDROPYRIDINE-3-CARBOXAMIDE (three-letter code: 35Q) (formula: $C_{29}H_{32}F_3N_5O_3$).
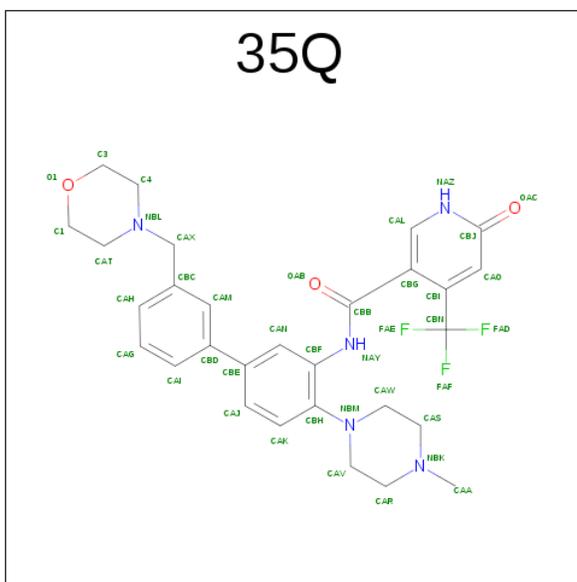
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 36	C 26	F 3	N 5	O 2	0	0
2	B	1	Total 36	C 26	F 3	N 5	O 2	0	0

- Molecule 3 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total 1	X 1	0	0
3	A	1	Total 1	X 1	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).
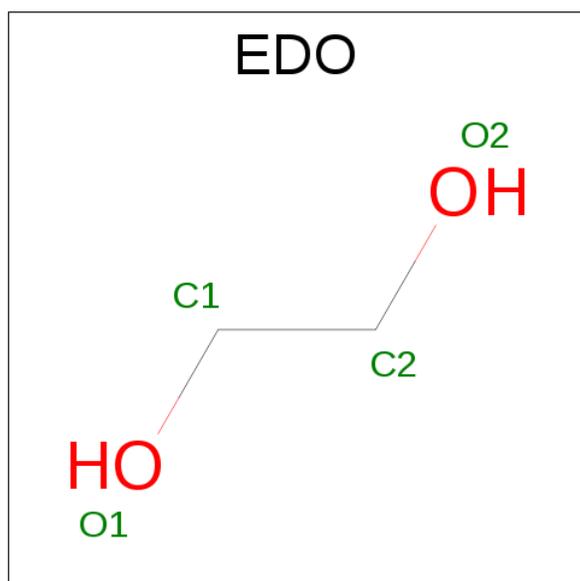
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total 4	C 2 O 2	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).
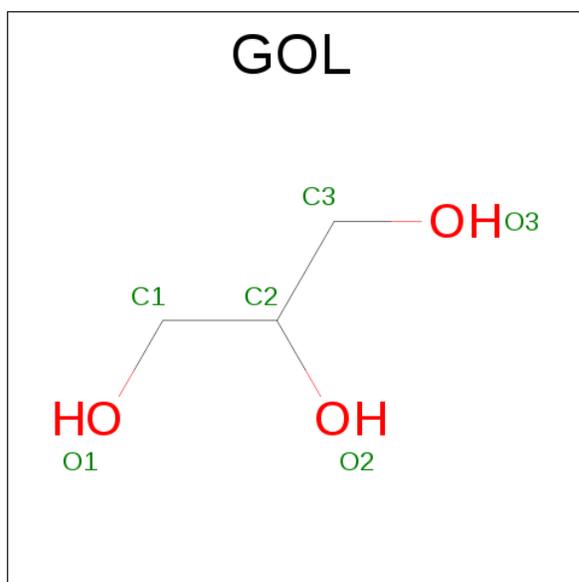
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total 6	C 3	O 3	0	0

- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	169	Total 173	O 173	0	4
6	B	157	Total 159	O 159	0	2

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

Note EDS failed to run properly.

• Molecule 1: WD repeat-containing protein 5

Chain A:

• Molecule 1: WD repeat-containing protein 5

Chain B:

4 Data and refinement statistics ⓘ

EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	46.84Å 56.63Å 64.69Å 71.45° 88.94° 65.43°	Depositor
Resolution (Å)	50.00 – 1.50	Depositor
% Data completeness (in resolution range)	94.5 (50.00-1.50)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	6.61 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.8.0071, COOT 0.7.2	Depositor
R, R_{free}	0.195 , 0.227	Depositor
Wilson B-factor (Å²)	13.2	Xtriage
Anisotropy	0.333	Xtriage
L-test for twinning[2]	$<|L|> = 0.48$, $<L^2> = 0.31$	Xtriage
Estimated twinning fraction	0.118 for h,h-k,-l	Xtriage
Total number of atoms	5115	wwPDB-VP
Average B, all atoms (Å²)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.67% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5   Model quality (i)

5.1   Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: UNX, GOL, EDO, 35Q

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$|Z|$>5	RMSZ	#$|Z|$>5
1	A	0.48	0/2414	0.65	0/3282
1	B	0.47	0/2435	0.66	0/3309
All	All	0.47	0/4849	0.65	0/6591

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2   Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2346	0	2280	4	0
1	B	2353	0	2308	12	0
2	A	36	0	24	0	0
2	B	36	0	24	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	B	4	0	6	0	0
5	B	6	0	8	3	0
6	A	173	0	0	0	0
6	B	159	0	0	1	0
All	All	5115	0	4650	16	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 2.

All (16) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:207:LYS:CD	1:B:243:TYR:O	2.48	0.62
1:B:263:PHE:CE1	5:B:403:GOL:H31	2.35	0.62
1:B:263:PHE:HE1	5:B:403:GOL:H31	1.66	0.60
1:A:209:LEU:HD12	1:A:229:ILE:HD11	1.88	0.55
1:B:91[A]:SER:HB2	6:B:569:HOH:O	2.11	0.50
1:B:97[A]:SER:OG	1:B:139:PRO:O	2.27	0.50
1:B:140[A]:GLN:H	1:B:140[A]:GLN:HE21	1.58	0.50
1:A:63:SER:HB3	1:A:93:VAL:HG13	1.97	0.46
1:B:140[B]:GLN:HA	1:B:140[B]:GLN:HE21	1.80	0.46
1:B:41:LEU:HB2	1:B:327:ILE:HB	1.98	0.45
1:A:41:LEU:HB2	1:A:327:ILE:HB	1.99	0.45
1:B:263:PHE:CE1	5:B:403:GOL:C3	3.01	0.43
1:A:303:VAL:HB	1:A:321:LEU:HD12	2.00	0.42
1:B:218:SER:HB2	1:B:261:CYS:HA	2.02	0.41
1:B:35:TYR:CG	1:B:332:SER:HB2	2.56	0.41
1:B:220:VAL:HG23	1:B:229[B]:ILE:HG23	2.03	0.40

There are no symmetry-related clashes.

5.3　Torsion angles ⓘ

5.3.1　Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	308/312 (99%)	295 (96%)	13 (4%)	0	100	100
1	B	307/312 (98%)	297 (97%)	10 (3%)	0	100	100
All	All	615/624 (99%)	592 (96%)	23 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	256/273 (94%)	256 (100%)	0	100	100
1	B	262/273 (96%)	259 (99%)	3 (1%)	78	56
All	All	518/546 (95%)	515 (99%)	3 (1%)	93	76

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	41	LEU
1	B	140[A]	GLN
1	B	140[B]	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 6 ligands modelled in this entry, 2 are unknown - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and

the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	#$|Z| > 2$	Counts	RMSZ	#$|Z| > 2$
2	35Q	A	401	-	37,39,44	1.76	9 (24%)	51,56,63	1.48	7 (13%)
2	35Q	B	401	-	37,39,44	1.70	7 (18%)	51,56,63	1.50	10 (19%)
4	EDO	B	402	-	3,3,3	0.49	0	2,2,2	0.32	0
5	GOL	B	403	-	5,5,5	0.10	0	5,5,5	0.64	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	35Q	A	401	-	-	0/24/35/44	0/4/4/5
2	35Q	B	401	-	-	0/24/35/44	0/4/4/5
4	EDO	B	402	-	-	0/1/1/1	0/0/0/0
5	GOL	B	403	-	-	0/4/4/4	0/0/0/0

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	401	35Q	CAX-CBC	-5.40	1.39	1.51
2	B	401	35Q	CAX-CBC	-4.58	1.41	1.51
2	A	401	35Q	CBN-CBI	-4.25	1.40	1.50
2	B	401	35Q	CBG-CBB	-4.23	1.41	1.50
2	B	401	35Q	CBN-CBI	-4.07	1.41	1.50
2	A	401	35Q	CBG-CBB	-3.42	1.43	1.50
2	B	401	35Q	FAF-CBN	-3.25	1.21	1.32
2	A	401	35Q	CBE-CBD	-2.57	1.42	1.49
2	B	401	35Q	CBF-NAY	-2.51	1.36	1.41
2	B	401	35Q	CBH-NBM	-2.34	1.36	1.41
2	A	401	35Q	CBH-NBM	-2.34	1.36	1.41
2	A	401	35Q	FAF-CBN	-2.18	1.25	1.32
2	A	401	35Q	CBF-NAY	-2.16	1.37	1.41
2	A	401	35Q	FAE-CBN	2.24	1.40	1.32
2	B	401	35Q	CBJ-NAZ	3.22	1.38	1.33
2	A	401	35Q	CBJ-NAZ	4.03	1.40	1.33

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	401	35Q	CAO-CBJ-NAZ	-4.58	118.21	124.79
2	B	401	35Q	CAO-CBJ-NAZ	-3.97	119.09	124.79
2	B	401	35Q	CAK-CBH-NBM	-3.30	117.09	122.28
2	B	401	35Q	CAW-CAS-NBK	-3.24	107.07	110.79
2	A	401	35Q	CAK-CBH-NBM	-2.57	118.25	122.28
2	B	401	35Q	FAE-CBN-CBI	-2.08	108.94	112.69
2	B	401	35Q	CAL-NAZ-CBJ	2.07	119.99	116.89
2	B	401	35Q	CBE-CAN-CBF	2.10	123.27	120.19
2	B	401	35Q	CAV-NBM-CBH	2.15	121.29	116.33
2	B	401	35Q	CBI-CBG-CBB	2.20	125.86	122.76
2	A	401	35Q	CAL-NAZ-CBJ	2.21	120.20	116.89
2	A	401	35Q	CAV-NBM-CBH	2.32	121.68	116.33
2	A	401	35Q	CBI-CBG-CBB	2.38	126.12	122.76
2	B	401	35Q	CBF-CBH-NBM	2.43	124.80	120.17
2	B	401	35Q	CAS-CAW-NBM	2.52	115.64	110.68
2	A	401	35Q	CAS-CAW-NBM	3.57	117.70	110.68
2	A	401	35Q	CAV-CAR-NBK	4.26	115.68	110.79

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	403	GOL	3	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

EDS failed to run properly - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

EDS failed to run properly - this section is therefore empty.

6.3 Carbohydrates ⓘ

EDS failed to run properly - this section is therefore empty.

6.4 Ligands ⓘ

EDS failed to run properly - this section is therefore empty.

6.5 Other polymers ⓘ

EDS failed to run properly - this section is therefore empty.