Full wwPDB X-ray Structure Validation Report (i)

Aug 29, 2023 – 07:26 pm BST

PDB ID	:	8Q1N
Title	:	Cyclic peptide binder of the WBM-site of WDR5
Authors	:	Schmeing, S.; Chang, J.Y.; t Hart, P.; Gasper, R.
Deposited on	:	2023-08-01
Resolution	:	1.84 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

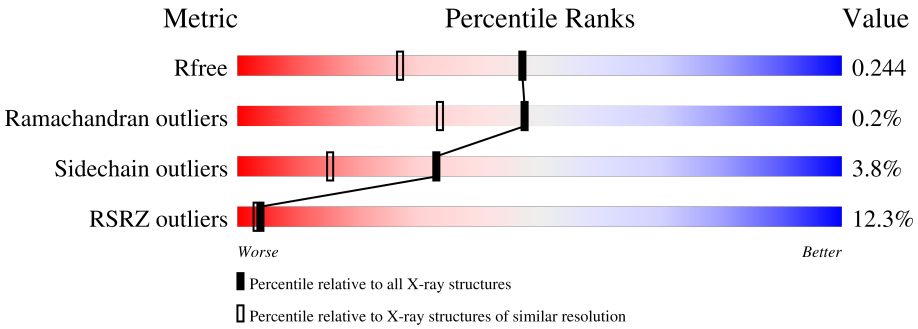
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 1.84 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	4003 (1.86-1.82)
Ramachandran outliers	138981	4185 (1.86-1.82)
Sidechain outliers	138945	4186 (1.86-1.82)
RSRZ outliers	127900	3957 (1.86-1.82)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. 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.

Mol	Chain	Length	Quality of chain
1	A	321	5% 96%
1	B	321	18% 95%
2	a	11	36% 100%
2	b	11	36% 91% 9%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5614 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	320	Total	C	N	O	S	0	10	0
			2525	1613	418	483	11			
1	B	317	Total	C	N	O	S	0	4	0
			2474	1581	410	472	11			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	335	GLY	-	expression tag	UNP P61964
A	336	THR	-	expression tag	UNP P61964
A	337	LEU	-	expression tag	UNP P61964
A	338	GLU	-	expression tag	UNP P61964
A	339	VAL	-	expression tag	UNP P61964
A	340	LEU	-	expression tag	UNP P61964
A	341	PHE	-	expression tag	UNP P61964
A	342	GLN	-	expression tag	UNP P61964
B	335	GLY	-	expression tag	UNP P61964
B	336	THR	-	expression tag	UNP P61964
B	337	LEU	-	expression tag	UNP P61964
B	338	GLU	-	expression tag	UNP P61964
B	339	VAL	-	expression tag	UNP P61964
B	340	LEU	-	expression tag	UNP P61964
B	341	PHE	-	expression tag	UNP P61964
B	342	GLN	-	expression tag	UNP P61964

- Molecule 2 is a protein called Cyclic peptide inhibitor.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	a	11	Total	C	N	O	0	0	0
			87	53	12	22			
2	b	11	Total	C	N	O	0	1	0
			93	57	12	24			

- Molecule 3 is water.

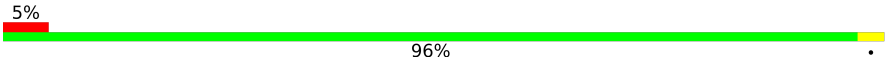
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	292	Total O 292 292	0	0
3	B	129	Total O 129 129	0	0
3	a	9	Total O 9 9	0	0
3	b	5	Total O 5 5	0	0

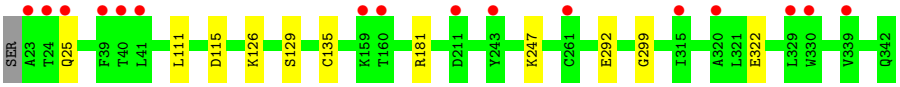
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

• Molecule 1: WD repeat-containing protein 5

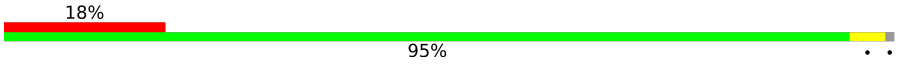
Chain A:

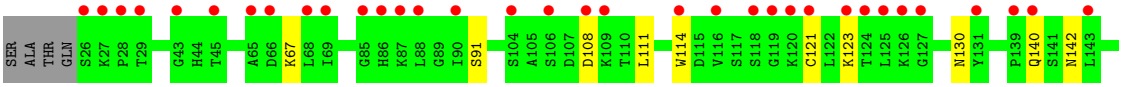


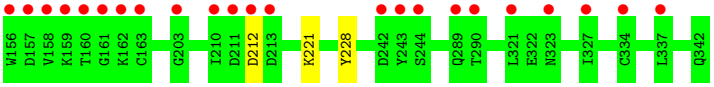


• Molecule 1: WD repeat-containing protein 5

Chain B:

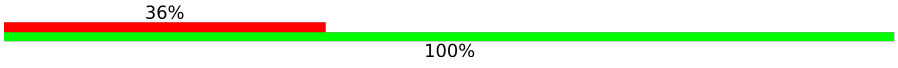






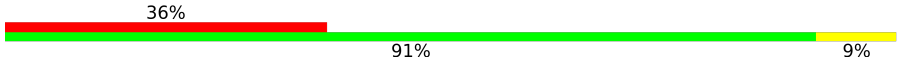
• Molecule 2: Cyclic peptide inhibitor

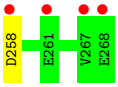
Chain a:



• Molecule 2: Cyclic peptide inhibitor

Chain b:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	82.17Å 82.17Å 201.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.98 – 1.84 43.96 – 1.84	Depositor EDS
% Data completeness (in resolution range)	96.6 (42.98-1.84) 96.6 (43.96-1.84)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.92 (at 1.84Å)	Xtriage
Refinement program	PHENIX (1.20.1-4487_4487: ???)	Depositor
R, R_{free}	0.198 , 0.243 0.198 , 0.244	Depositor DCC
R_{free} test set	3030 reflections (5.17%)	wwPDB-VP
Wilson B-factor (Å^2)	39.4	Xtriage
Anisotropy	0.141	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 50.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	5614	wwPDB-VP
Average B, all atoms (Å^2)	56.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.59% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: DAB

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	0.57	1/2611 (0.0%)	0.73	0/3542
1	B	0.50	0/2542	0.71	0/3450
2	a	0.48	0/78	0.68	0/103
2	b	0.36	0/87	0.53	0/115
All	All	0.53	1/5318 (0.0%)	0.72	0/7210

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	135	CYS	CB-SG	5.47	1.91	1.82

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

Due to software issues we are unable to calculate clashes - this section is therefore empty.

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	328/321 (102%)	314 (96%)	13 (4%)	1 (0%)	41 27
1	B	319/321 (99%)	300 (94%)	19 (6%)	0	100 100
2	a	8/11 (73%)	8 (100%)	0	0	100 100
2	b	9/11 (82%)	9 (100%)	0	0	100 100
All	All	664/664 (100%)	631 (95%)	32 (5%)	1 (0%)	47 33

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	299	GLY

5.3.2 Protein sidechains (i)

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	290/281 (103%)	280 (97%)	10 (3%)	37 19
1	B	282/281 (100%)	269 (95%)	13 (5%)	27 10
2	a	10/10 (100%)	10 (100%)	0	100 100
2	b	11/10 (110%)	10 (91%)	1 (9%)	9 1
All	All	593/582 (102%)	569 (96%)	24 (4%)	33 14

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25	GLN
1	A	111	LEU
1	A	115	ASP
1	A	126	LYS
1	A	129[A]	SER
1	A	129[B]	SER
1	A	181[A]	ARG
1	A	181[B]	ARG
1	A	247	LYS
1	A	322	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	67	LYS
1	B	91	SER
1	B	108	ASP
1	B	111	LEU
1	B	114	TRP
1	B	121	CYS
1	B	123	LYS
1	B	130	ASN
1	B	140	GLN
1	B	142	ASN
1	B	212	ASP
1	B	221	LYS
1	B	228	TYR
2	b	258	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	100	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

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	DAB	a	266	2	5,6,7	0.90	0	1,6,8	0.23	0
2	DAB	b	266	2	5,6,7	0.91	0	1,6,8	0.10	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	DAB	a	266	2	-	1/4/5/7	-
2	DAB	b	266	2	-	1/4/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	a	266	DAB	CA-CB-CG-ND
2	b	266	DAB	CA-CB-CG-ND

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

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

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	320/321 (99%)	0.42	16 (5%) 28 26	32, 44, 63, 132	0
1	B	317/321 (98%)	1.04	57 (17%) 1 1	36, 61, 94, 131	0
2	a	10/11 (90%)	2.35	4 (40%) 0 0	42, 86, 129, 139	0
2	b	10/11 (90%)	2.15	4 (40%) 0 0	63, 92, 116, 117	0
All	All	657/664 (98%)	0.77	81 (12%) 4 3	32, 51, 91, 139	0

All (81) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	a	258	ASP	8.8
1	B	158	VAL	8.4
1	B	119	GLY	7.8
1	B	162	LYS	7.7
2	b	268	GLU	7.3
1	A	25	GLN	7.2
1	A	24	THR	6.1
1	B	161	GLY	6.1
1	B	244	SER	5.9
1	B	121	CYS	5.7
1	B	85	GLY	5.4
1	B	243	TYR	5.3
1	B	125	LEU	5.1
1	B	212	ASP	4.9
2	a	259	GLU	4.5
1	B	87	LYS	4.4
1	B	45	THR	4.3
1	B	29	THR	4.3
1	B	68	LEU	4.1
1	B	28	PRO	4.1
1	B	163	CYS	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	123	LYS	3.9
1	B	211	ASP	3.9
2	a	260	GLU	3.8
1	B	108	ASP	3.7
1	B	120	LYS	3.7
1	B	131	TYR	3.7
1	B	114	TRP	3.6
1	B	160	THR	3.5
1	B	26	SER	3.4
2	b	267	VAL	3.4
1	B	210	ILE	3.1
2	b	258	ASP	3.1
1	B	157	ASP	3.1
2	b	261	GLU	3.1
1	A	23	ALA	3.1
2	a	268	GLU	3.0
1	B	124	THR	3.0
1	A	315	ILE	3.0
1	B	66	ASP	2.9
1	A	329	LEU	2.9
1	B	118	SER	2.9
1	A	159	LYS	2.9
1	B	159	LYS	2.9
1	B	106	SER	2.8
1	B	334	CYS	2.8
1	B	203	GLY	2.8
1	B	116	VAL	2.8
1	B	213	ASP	2.8
1	B	323	ASN	2.7
1	A	320	ALA	2.7
1	B	337	LEU	2.7
1	B	88	LEU	2.6
1	B	43	GLY	2.6
1	B	104	SER	2.6
1	B	156	TRP	2.5
1	B	109	LYS	2.4
1	B	321	LEU	2.4
1	A	261[A]	CYS	2.4
1	A	330	TRP	2.4
1	B	86	HIS	2.4
1	B	327	ILE	2.4
1	A	40[A]	THR	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	90	ILE	2.3
1	B	127	GLY	2.3
1	B	242	ASP	2.3
1	B	69	ILE	2.3
1	B	126	LYS	2.3
1	A	39	PHE	2.2
1	B	289	GLN	2.2
1	A	211	ASP	2.2
1	B	143	LEU	2.2
1	A	160	THR	2.2
1	B	290	THR	2.2
1	A	243	TYR	2.1
1	B	140	GLN	2.1
1	B	65	ALA	2.1
1	B	27	LYS	2.1
1	B	139	PRO	2.1
1	A	339	VAL	2.0
1	A	41	LEU	2.0

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	DAB	b	266	7/8	0.83	0.21	55,73,87,92	0
2	DAB	a	266	7/8	0.95	0.10	49,51,61,66	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.