



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Sep 17, 2017 – 11:42 PM EDT

PDB ID : 5KMG
EMDB ID: : EMD-8266
Title : Near-atomic cryo-EM structure of PRC1 bound to the microtubule
Authors : Kellogg, E.H.; Howes, S.; Ti, S.-C.; Ramirez-Aportela, E.; Kapoor, T.M.; Chacon, P.; Nogales, E.
Deposited on : unknown
Resolution : 3.50 Å(reported)
Based on PDB ID : 3NRX, 4L6Y

This is a wwPDB/EMDataBank EM Map/Model Validation Summary Report
for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
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) : rb-20029824

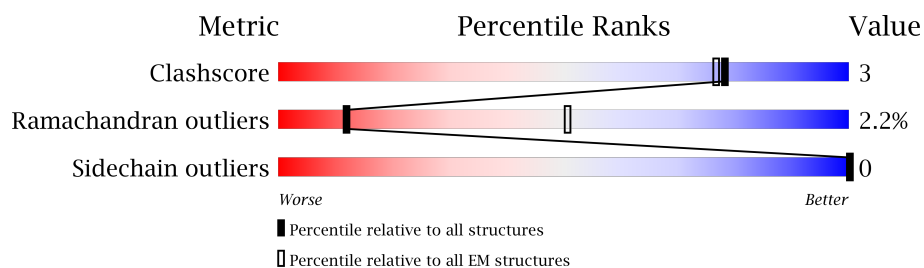
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 3.50 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	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	441	92% 6% .
2	B	431	94% 5% .
3	P	128	93% 7%

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 7991 atoms, of which 20 are hydrogens and 0 are deuteriums.

In the tables below, 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 Tubulin alpha-1B chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	431	Total	C	N	O	S	0	0
			3383	2144	574	643	22		

- Molecule 2 is a protein called Tubulin beta chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	431	Total	C	N	O	S	0	0
			3385	2124	580	655	26		

- Molecule 3 is a protein called Protein regulator of cytokinesis 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	P	128	Total	C	N	O	S	0	0
			1104	700	197	202	5		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
P	337	GLY	-	expression tag	UNP O43663
P	338	ALA	-	expression tag	UNP O43663
P	339	ALA	-	expression tag	UNP O43663
P	340	ALA	-	expression tag	UNP O43663

- Molecule 4 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).

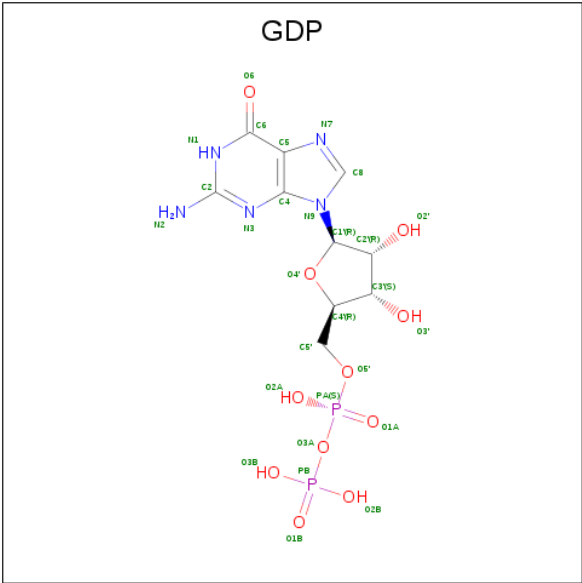


Mol	Chain	Residues	Atoms						AltConf
4	A	1	Total	C	H	N	O	P	0
			42	10	10	5	14	3	

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

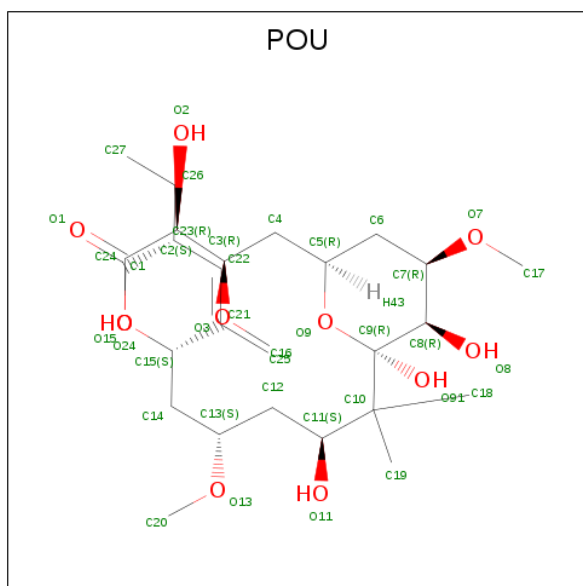
Mol	Chain	Residues	Atoms		AltConf
5	A	1	Total	Mg	0
			1	1	

- Molecule 6 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: C₁₀H₁₅N₅O₁₁P₂).



Mol	Chain	Residues	Atoms					AltConf	
6	B	1	Total	C	H	N	O	P	0
			38	10	10	5	11	2	

- Molecule 7 is Peloruside A (three-letter code: POA) (formula: $C_{27}H_{48}O_{11}$).

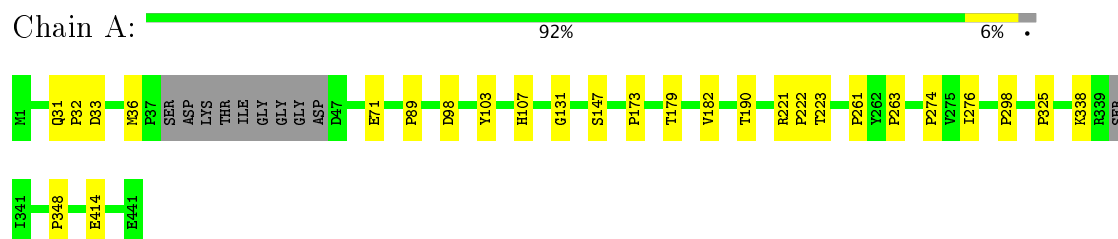


Mol	Chain	Residues	Atoms			AltConf
7	B	1	Total	C	O	0
			38	27	11	

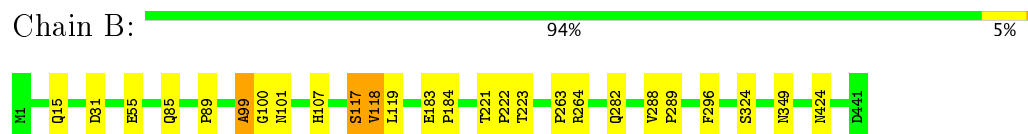
3 Residue-property plots [i](#)

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. 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. 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.

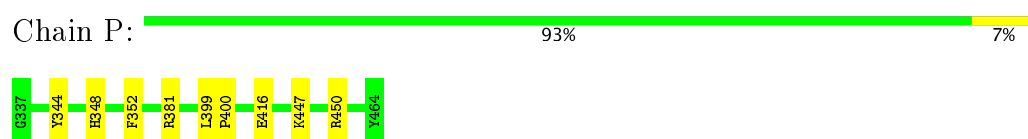
- Molecule 1: Tubulin alpha-1B chain



- Molecule 2: Tubulin beta chain



- Molecule 3: Protein regulator of cytokinesis 1



4 Experimental information

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	HELICAL, twist=-25.76°, rise=8.67 Å, axial sym=C1	Depositor
Number of segments used	17069	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{Å}^2$)	8	Depositor
Minimum defocus (nm)	1100	Depositor
Maximum defocus (nm)	3800	Depositor
Magnification	27500	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

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: GDP, GTP, MG, POU

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 >2	RMSZ	# Z >2
1	A	1.05	0/3459	0.61	0/4693
2	B	1.02	0/3460	0.57	0/4687
3	P	0.97	0/1126	0.55	0/1499
All	All	1.03	0/8045	0.59	0/10879

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	3383	0	3295	14	0
2	B	3385	0	3262	25	0
3	P	1104	0	1096	15	0
4	A	32	10	12	0	0
5	A	1	0	0	0	0
6	B	28	10	12	0	0
7	B	38	0	48	11	0
All	All	7971	20	7725	52	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 3.

The worst 5 of 52 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)
3:P:344:TYR:CZ	3:P:416:GLU:HB3	1.41	1.50
3:P:344:TYR:OH	3:P:416:GLU:HB3	1.37	1.20
3:P:344:TYR:CZ	3:P:416:GLU:CB	2.36	1.08
2:B:424:ASN:OD1	3:P:381:ARG:NH2	1.91	1.03
2:B:296:PHE:HD2	7:B:502:POU:H24	1.24	0.99

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

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	425/441 (96%)	384 (90%)	27 (6%)	14 (3%)	4	35
2	B	429/431 (100%)	393 (92%)	28 (6%)	8 (2%)	9	47
3	P	126/128 (98%)	124 (98%)	2 (2%)	0	100	100
All	All	980/1000 (98%)	901 (92%)	57 (6%)	22 (2%)	12	44

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	179	THR
1	A	338	LYS
2	B	100	GLY
2	B	118	VAL
1	A	32	PRO

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 PDB entries followed by that with respect to all EM

entries.

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	365/372 (98%)	365 (100%)	0	100	100
2	B	370/370 (100%)	370 (100%)	0	100	100
3	P	114/114 (100%)	114 (100%)	0	100	100
All	All	849/856 (99%)	849 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	8	GLN
2	B	11	GLN
2	B	15	GLN

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 4 ligands modelled in this entry, 1 is monoatomic - leaving 3 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	$\# Z > 2$
4	GTP	A	501	5	27,34,34	1.45	4 (14%)	24,54,54	1.45	3 (12%)
6	GDP	B	501	-	25,30,30	1.68	6 (24%)	23,47,47	1.44	2 (8%)
7	POU	B	502	-	35,39,39	2.31	7 (20%)	42,57,57	2.09	11 (26%)

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
4	GTP	A	501	5	-	0/18/38/38	0/3/3/3
6	GDP	B	501	-	-	0/12/32/32	0/3/3/3
7	POU	B	502	-	-	1/54/76/76	0/0/2/2

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	502	POU	C2-C1	-5.09	1.39	1.52
4	A	501	GTP	C6-N1	-3.18	1.32	1.36
6	B	501	GDP	C6-N1	-2.93	1.32	1.36
6	B	501	GDP	C2-N1	-2.57	1.32	1.36
4	A	501	GTP	C2-N1	-2.47	1.33	1.36

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	B	502	POU	C13-C12-C11	-7.70	104.60	114.42
6	B	501	GDP	C1'-N9-C4	-4.49	118.88	126.64
7	B	502	POU	C3-C4-C5	-3.01	107.71	114.79
7	B	502	POU	C26-C23-C24	-2.80	105.76	111.80
4	A	501	GTP	C4-C5-N7	-2.80	106.71	109.41

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	B	502	POU	C23-C22-C21-C15

There are no ring outliers.

1 monomer is involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	502	POU	11	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.