WORLDWIDE PDB PROTEIN DATA BANK

EMDataBank
Unified Data Resource for 3DEM

Full wwPDB/EMDataBank EM Map/Model Validation Report (i)

Sep 17, 2017 – 01:20 AM EDT

PDB ID : 5IJ9
EMDB ID: : EMD-8095
Title : Cryo EM density of microtubule assembled from human TUBB3-D417H mutant
Authors : Ti, S.C.; Pamula, M.C.; Howes, S.C.; Duellberg, C.; Cade, N.I.; Kleiner, R.E.; Forth, S.; Surrey, T.; Nogales, E.; Kapoor, T.M.
Deposited on : unknown
Resolution : 3.70 Å(reported)

This is a Full wwPDB/EMDataBank EM Map/Model Validation 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
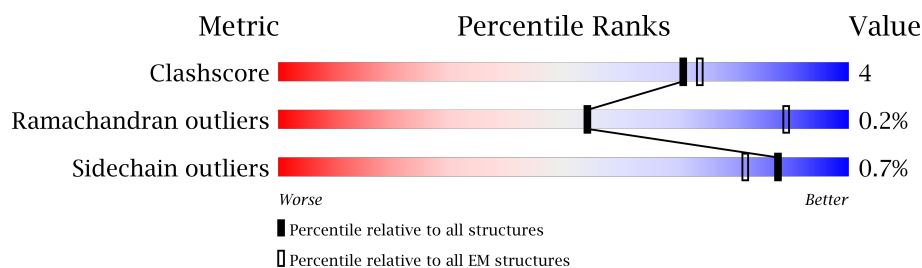with specific help available everywhere you see the (i) symbol.

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 3.70 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		4
Ramachandran outliers		0.2%
Sidechain outliers		0.7%

Worse   *Better*

▮ Percentile relative to all structures
▯ Percentile relative to all EM structures

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	437	86% 12% ·
2	B	426	89% 10%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 13309 atoms, of which 6526 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	430	Total 6653	C 2136	H 3283	N 573	O 639	S 22	0	0

- Molecule 2 is a protein called Tubulin beta-3 chain.

Mol	Chain	Residues	Atoms						AltConf	Trace
2	B	426	Total 6595	C 2110	H 3243	N 576	O 640	S 26	0	0

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	417	HIS	ASP	engineered mutation	UNP Q13509

- Molecule 3 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).
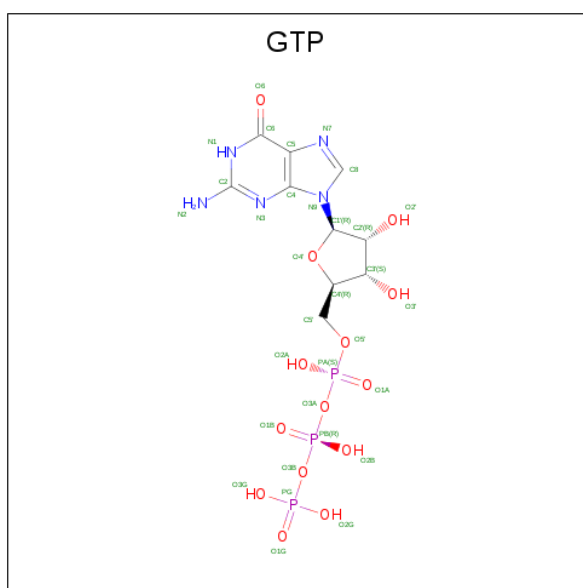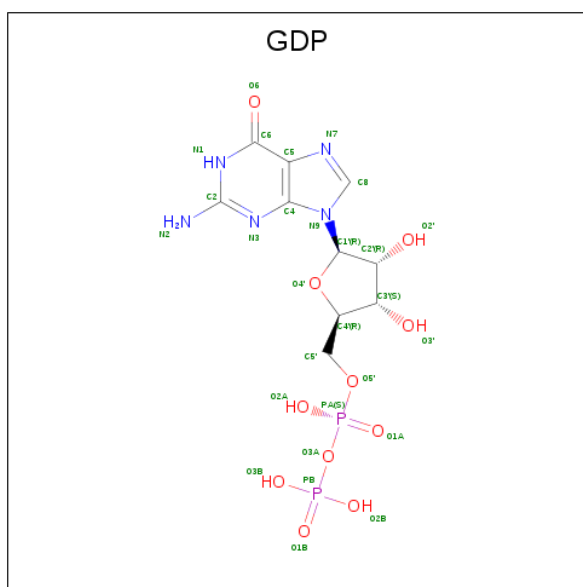
Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
3	A	1	32	10	5	14	3	0

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 5 is G U A N O S I N E - 5' - D I P H O S P H A T E (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).

| Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
5	B	1	28	10	5	11	2	0

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. 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
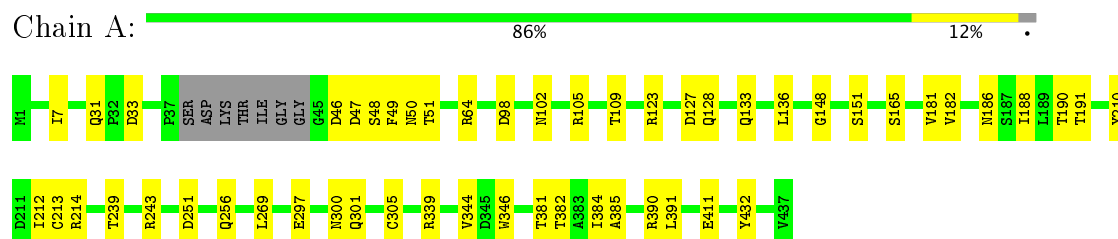
Chain A:

- Molecule 2: Tubulin beta-3 chain

Chain B:
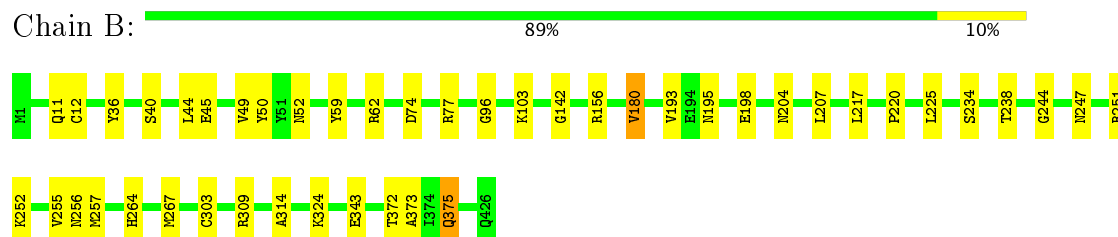
4   Experimental information (i)

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	HELICAL, twist=-27.68°, rise=9.48 Å, axial sym=C1	Depositor
Number of segments used	26959	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 $(e^-/Å^2)$	27.6	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	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

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	0.24	0/3447	0.39	0/4679
2	B	0.23	0/3428	0.39	0/4642
All	All	0.23	0/6875	0.39	0/9321

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	3370	3283	3284	30	0
2	B	3352	3243	3245	27	0
3	A	32	0	12	0	0
4	A	1	0	0	0	0
5	B	28	0	12	3	0
All	All	6783	6526	6553	55	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 4.

All (55) 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)
2:B:11:GLN:NE2	5:B:501:GDP:O1B	2.19	0.76
2:B:156:ARG:NE	2:B:195:ASN:O	2.22	0.72
1:A:48:SER:O	1:A:51:THR:OG1	2.06	0.71
2:B:96:GLY:O	2:B:103:LYS:NZ	2.27	0.64
1:A:102:ASN:ND2	1:A:411:GLU:OE1	2.30	0.63
2:B:52:ASN:OD1	2:B:62:ARG:NH2	2.30	0.63
2:B:49:VAL:O	2:B:62:ARG:NH1	2.32	0.63
1:A:133:GLN:NE2	1:A:251:ASP:OD2	2.32	0.63
2:B:142:GLY:N	5:B:501:GDP:O2B	2.34	0.60
2:B:244:GLY:O	2:B:247:ASN:ND2	2.36	0.58
2:B:204:ASN:ND2	5:B:501:GDP:O2'	2.37	0.57
1:A:186:ASN:O	1:A:190:THR:OG1	2.09	0.56
2:B:74:ASP:OD1	2:B:77:ARG:NH2	2.39	0.56
1:A:47:ASP:O	1:A:50:ASN:N	2.39	0.55
2:B:252:LYS:O	2:B:256:ASN:ND2	2.37	0.55
1:A:239:THR:O	1:A:243:ARG:N	2.39	0.55
1:A:123:ARG:NH1	1:A:127:ASP:OD2	2.42	0.53
1:A:48:SER:HA	1:A:51:THR:HG23	1.91	0.53
1:A:7:ILE:N	1:A:136:LEU:O	2.42	0.50
2:B:49:VAL:HG13	2:B:50:TYR:H	1.77	0.50
2:B:309:ARG:NH1	2:B:343:GLU:OE1	2.46	0.49
2:B:207:LEU:HB3	2:B:225:LEU:HD22	1.95	0.49
2:B:49:VAL:HG13	2:B:50:TYR:N	2.28	0.48
1:A:212:ILE:HD11	1:A:301:GLN:O	2.13	0.48
2:B:303:CYS:HB2	2:B:373:ALA:HB1	1.95	0.48
1:A:148:GLY:O	1:A:151:SER:OG	2.27	0.47
1:A:31:GLN:NE2	1:A:33:ASP:OD2	2.48	0.46
1:A:382:THR:O	1:A:385:ALA:HB3	2.16	0.46
2:B:36:TYR:OH	2:B:40:SER:O	2.34	0.45
1:A:269:LEU:HD21	1:A:384:ILE:HD13	1.98	0.45
1:A:297:GLU:OE1	1:A:300:ASN:ND2	2.48	0.45
2:B:198:GLU:HA	2:B:264:HIS:HB2	1.98	0.45
1:A:381:THR:O	1:A:384:ILE:HG22	2.17	0.45
2:B:180:VAL:HG12	2:B:180:VAL:O	2.15	0.45
1:A:181:VAL:HG23	1:A:182:VAL:HG13	1.98	0.44
1:A:385:ALA:HB2	1:A:432:TYR:CG	2.53	0.44
1:A:46:ASP:O	1:A:49:PHE:N	2.36	0.44
2:B:234:SER:O	2:B:238:THR:N	2.52	0.43
1:A:165:SER:OG	1:A:256:GLN:NE2	2.52	0.42
2:B:372:THR:O	2:B:375:GLN:N	2.52	0.42
1:A:188:ILE:HA	1:A:191:THR:HG22	2.00	0.42
2:B:217:LEU:HD13	2:B:220:PRO:HG3	2.01	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:257:MET:SD	2:B:314:ALA:HB2	2.60	0.42
1:A:105:ARG:O	1:A:109:THR:N	2.53	0.42
1:A:98:ASP:O	2:B:251:ARG:NH2	2.48	0.41
1:A:47:ASP:C	1:A:49:PHE:N	2.73	0.41
2:B:44:LEU:HD22	2:B:59:TYR:CZ	2.55	0.41
2:B:251:ARG:O	2:B:255:VAL:HG23	2.20	0.41
2:B:267:MET:HE1	2:B:373:ALA:HB3	2.02	0.41
1:A:51:THR:O	1:A:64:ARG:NH1	2.54	0.41
1:A:390:ARG:HG3	1:A:391:LEU:N	2.37	0.40
1:A:210:TYR:CG	2:B:324:LYS:HD3	2.56	0.40
1:A:339:ARG:HA	1:A:339:ARG:NE	2.35	0.40
1:A:344:VAL:HG11	1:A:346:TRP:NE1	2.37	0.40
1:A:213:CYS:SG	1:A:214:ARG:N	2.94	0.40

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	426/437 (98%)	402 (94%)	24 (6%)	0	100	100
2	B	424/426 (100%)	403 (95%)	19 (4%)	2 (0%)	32	73
All	All	850/863 (98%)	805 (95%)	43 (5%)	2 (0%)	54	85

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	180	VAL
2	B	45	GLU

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	363/368 (99%)	361 (99%)	2 (1%)	89	95
2	B	367/367 (100%)	364 (99%)	3 (1%)	85	93
All	All	730/735 (99%)	725 (99%)	5 (1%)	87	94

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	128	GLN
1	A	305	CYS
2	B	12	CYS
2	B	193	VAL
2	B	375	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	256	GLN
2	B	35	ASN
2	B	131	GLN
2	B	204	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)

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

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	#$|Z| > 2$
3	GTP	A	501	4	27,34,34	0.93	1 (3%)	27,54,54	1.67	4 (14%)
5	GDP	B	501	-	25,30,30	1.15	2 (8%)	26,47,47	1.97	6 (23%)

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
3	GTP	A	501	4	-	0/18/38/38	0/3/3/3
5	GDP	B	501	-	-	0/12/32/32	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	501	GTP	C6-N1	2.95	1.38	1.33
5	B	501	GDP	C5-C4	3.06	1.47	1.40
5	B	501	GDP	C6-C5	3.84	1.48	1.41

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	501	GTP	N3-C2-N1	-4.99	120.17	127.46
5	B	501	GDP	C5-C6-N1	-3.85	118.00	123.48
5	B	501	GDP	C6-C5-C4	-3.66	117.20	120.84
5	B	501	GDP	N3-C2-N1	-3.16	122.85	127.46

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	B	501	GDP	C4-C5-N7	-2.84	106.67	109.41
3	A	501	GTP	C5-C6-N1	-2.65	119.71	123.48
3	A	501	GTP	C6-N1-C2	2.60	119.80	116.06
5	B	501	GDP	C6-N1-C2	4.36	122.33	116.06
3	A	501	GTP	C2-N3-C4	4.40	120.29	115.16
5	B	501	GDP	C2-N3-C4	4.94	120.93	115.16

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	501	GDP	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.