wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:53 PM BST

PDB ID	:	3JAW
EMDB ID:	:	EMD-6354
Title	:	Atomic model of a microtubule seam based on a cryo-EM reconstruction of the EB3-bound microtubule (merged dataset containing tubulin bound to GTPgammaS, GMPCPP, and GDP)
Authors	:	Zhang, R.; Nogales, E.
Deposited on	:	2015-06-30
Resolution	:	3.90 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could
stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/EMValidationReportHelp

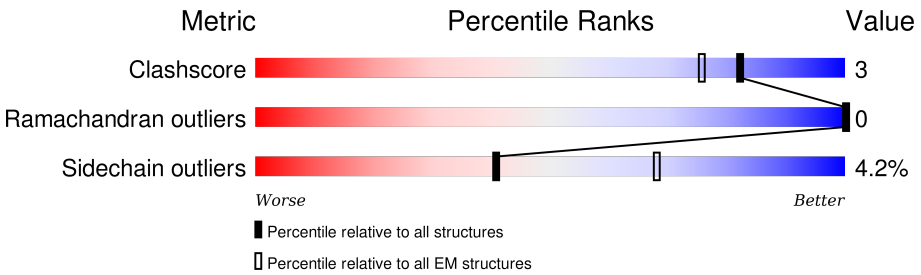
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 3.90 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	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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	451	83% 13% •
1	C	451	88% 7% •
2	B	445	85% 11% •
2	D	445	85% 11% • •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 13632 atoms, of which 0 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.

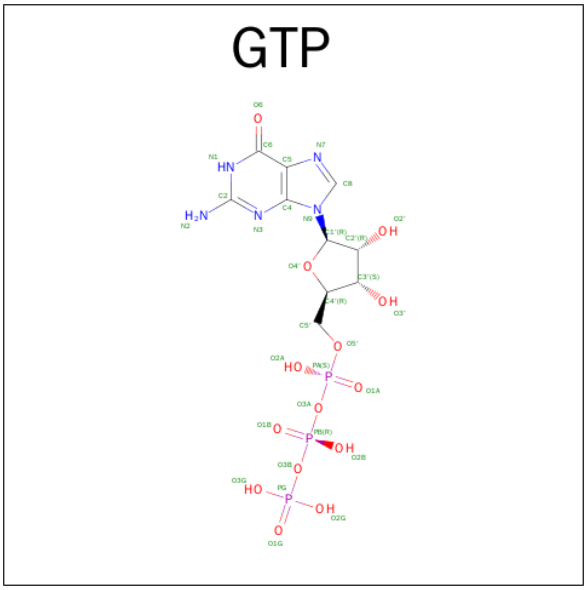
- Molecule 1 is a protein called Tubulin alpha-1B chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	432	3383	2144	575	643	21	0	0
1	C	432	3383	2144	575	643	21	0	0

- Molecule 2 is a protein called Tubulin beta chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	429	3368	2115	578	650	25	0	0
2	D	429	3368	2115	578	650	25	0	0

- Molecule 3 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).

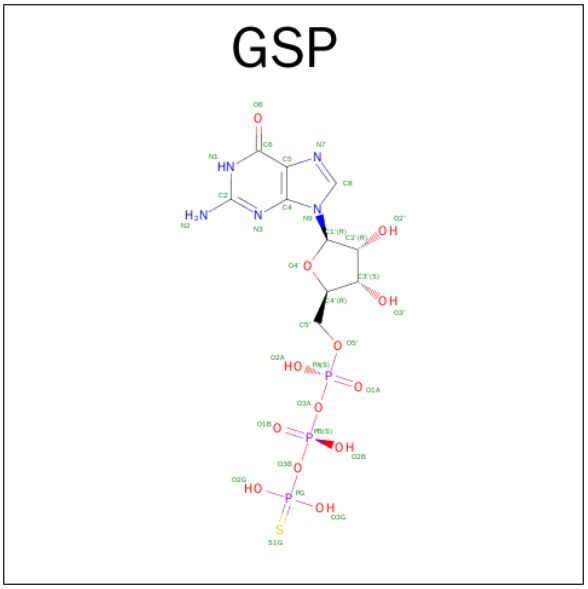


Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
3	A	1	32	10	5	14	3	0
			Total	C	N	O	P	
3	C	1	32	10	5	14	3	0

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

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

- Molecule 5 is 5'-GUANOSINE-DIPHOSPHATE-MONOTHIOPHOSPHATE (three-letter code: GSP) (formula: $C_{10}H_{16}N_5O_{13}P_3S$).



Mol	Chain	Residues	Atoms						AltConf
			Total	C	N	O	P	S	
5	B	1	32	10	5	13	3	1	0
			Total	C	N	O	P	S	
5	D	1	32	10	5	13	3	1	0

3 Residue-property plots (i)

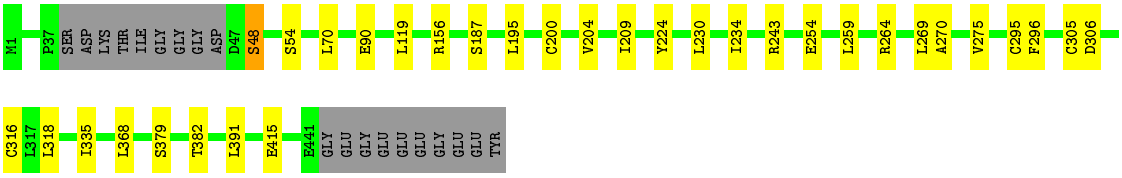
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric 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:

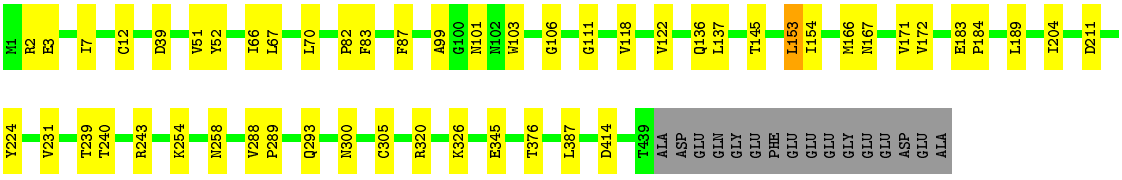
- Molecule 1: Tubulin alpha-1B chain

Chain C:



- Molecule 2: Tubulin beta chain

Chain B:



- Molecule 2: Tubulin beta chain

Chain D:

V195
I204
T221
Y224
G225
D226
S232
T239
R243
P263
H266
V288
P289
K299
M325
N350
C356
I384
L428
Y432
T439
ALA
ASP
GLU
GLN
GLY
GLU
PHE
GLU
GLU
GLU
GLY
GLU
GLU
ASP
GLU
ALA

4 Experimental information (i)

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	79702	Depositor
Resolution determination method	FSC 0.143	Depositor
CTF correction method	CTFFIND4, each particle	Depositor
Microscope	FEI TITAN	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	27.6	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	3500	Depositor
Magnification	27500	Depositor
Image detector	GATAN K2 (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: GTP, GSP, 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	0.44	0/3460	0.70	0/4698
1	C	0.44	0/3460	0.70	0/4698
2	B	0.43	0/3443	0.70	0/4666
2	D	0.44	0/3443	0.69	0/4666
All	All	0.44	0/13806	0.70	0/18728

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	3289	29	0
1	C	3383	0	3289	12	0
2	B	3368	0	3246	22	0
2	D	3368	0	3246	22	0
3	A	32	0	12	0	0
3	C	32	0	12	0	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
5	B	32	0	12	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	D	32	0	12	0	0
All	All	13632	0	13118	84	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 84 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:A:269:LEU:HD23	1:A:384:ILE:HD11	1.71	0.72
2:D:66:ILE:HD13	2:D:122:VAL:HG22	1.73	0.71
2:B:204:ILE:HD13	2:B:231:VAL:HG22	1.72	0.70
2:B:66:ILE:HD13	2:B:122:VAL:HG22	1.76	0.66
2:D:51:VAL:HG11	2:D:243:ARG:HG2	1.77	0.66

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	428/451 (95%)	416 (97%)	12 (3%)	0	100	100
1	C	428/451 (95%)	411 (96%)	17 (4%)	0	100	100
2	B	427/445 (96%)	417 (98%)	10 (2%)	0	100	100
2	D	427/445 (96%)	417 (98%)	10 (2%)	0	100	100
All	All	1710/1792 (95%)	1661 (97%)	49 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	364/379 (96%)	351 (96%)	13 (4%)	42	76
1	C	364/379 (96%)	348 (96%)	16 (4%)	35	71
2	B	368/381 (97%)	352 (96%)	16 (4%)	35	72
2	D	368/381 (97%)	352 (96%)	16 (4%)	35	72
All	All	1464/1520 (96%)	1403 (96%)	61 (4%)	41	72

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	376	THR
1	C	224	TYR
2	D	226	ASP
2	B	414	ASP
1	C	54	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	88	HIS
1	C	107	HIS
1	C	293	ASN
1	C	15	GLN
1	C	358	GLN

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 6 ligands modelled in this entry, 2 are monoatomic - 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	$\#\|Z\| > 2$
3	GTP	A	501	4	26,34,34	1.08	2 (7%)	29,54,54	1.74	5 (17%)
5	GSP	B	501	-	26,34,34	1.43	4 (15%)	27,54,54	2.09	7 (25%)
3	GTP	C	501	4	26,34,34	1.18	2 (7%)	29,54,54	1.92	5 (17%)
5	GSP	D	501	-	26,34,34	1.35	5 (19%)	27,54,54	2.11	7 (25%)

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	GSP	B	501	-	-	0/17/38/38	0/3/3/3
3	GTP	C	501	4	-	0/18/38/38	0/3/3/3
5	GSP	D	501	-	-	0/17/38/38	0/3/3/3

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	D	501	GSP	PG-O3G	2.17	1.61	1.55
5	B	501	GSP	PG-O2G	2.29	1.61	1.55
5	D	501	GSP	PG-O2G	2.34	1.61	1.55
5	B	501	GSP	PG-S1G	2.66	1.95	1.90
3	A	501	GTP	C5-C4	2.68	1.46	1.40

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	D	501	GSP	PB-O3B-PG	-4.30	117.10	132.71
3	C	501	GTP	C1'-N9-C4	-4.22	122.10	126.81
5	D	501	GSP	C5-C6-N1	-4.18	118.05	123.52
5	B	501	GSP	C5-C6-N1	-4.08	118.19	123.52
3	C	501	GTP	C5-C6-N1	-3.94	118.38	123.52

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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.