



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Jan 25, 2017 – 11:50 AM EST

PDB ID : 5SYF
EMDB ID: : EMD-8322
Title : High-resolution cryo-EM reconstruction of Taxol-stabilized microtubule
Authors : Kellogg, E.H.; Nogales, E.
Deposited on : 2016-08-11
Resolution : 3.50 Å(reported)

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 : 1.7.1 (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) : rb-20028442

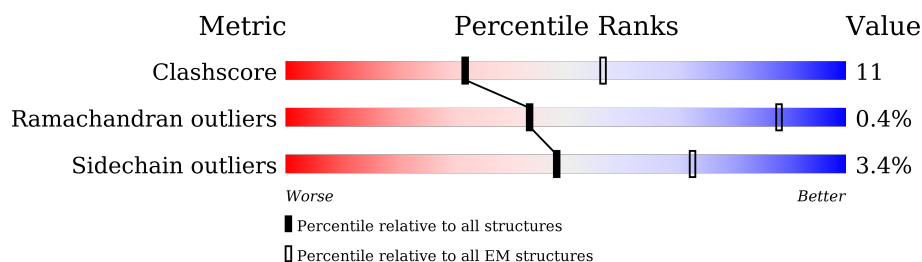
1 Overall quality at a glance

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	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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	437	
2	B	426	

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 6818 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.

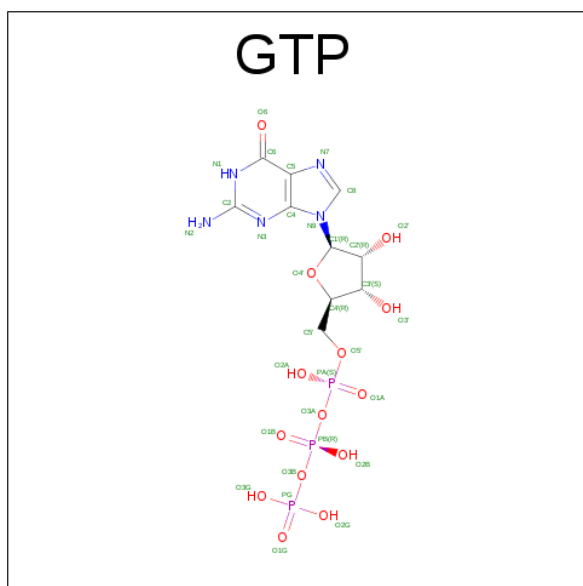
- Molecule 1 is a protein called Tubulin alpha chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	A	426	3343	2121	569	631	22	0	0

- Molecule 2 is a protein called Tubulin beta chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	426	3352	2106	575	645	26	0	0

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

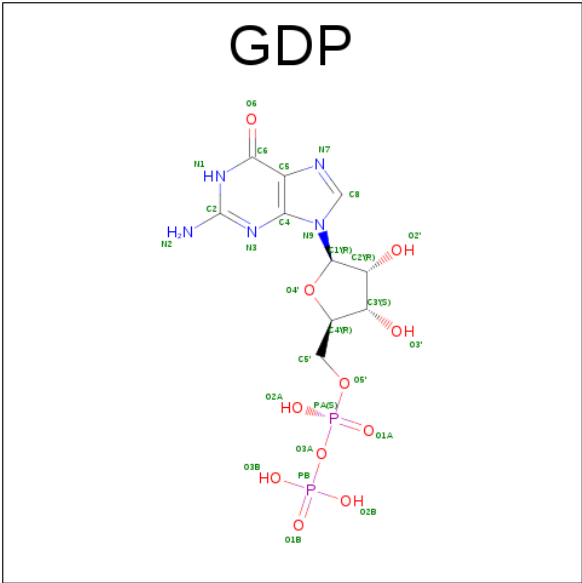


Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
3	A	1	32	10	5	14	3	0

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

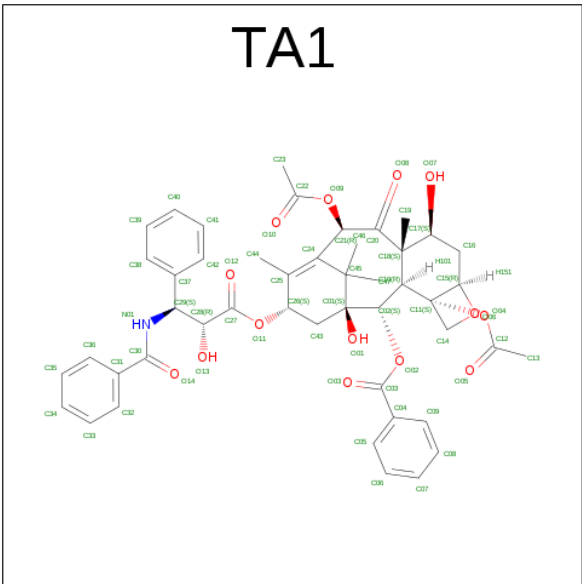
Mol	Chain	Residues	Atoms		AltConf
4	A	1	Total	Mg	0
			1	1	

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



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

- Molecule 6 is TAXOL (three-letter code: TA1) (formula: C₄₇H₅₁NO₁₄).



Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
6	B	1	62	47	1	14	0

- Molecule 1: Tubulin alpha chain

	P184	M1
G436 V437	I209	R2
	Y210	E3
	D211	O4
	I212	G10
	L217	Q11
	R221	V14
	Y224	O15
	L227	I16
	L230	P36
	V235	PRO
A240	S241	SER
L242	S241	ASP
L248	L242	LYS
L252	L248	THR
L269	L252	ILE
A270	L269	GLY
T271	A270	GLY
V275	T271	ASP
V303	V275	ASP
Y319	V303	S48
V344	Y319	F49
D345	V344	N50
W346	D345	T51
G354	W346	T73
I355	G354	H88
R373	I355	G95
M377	R373	D98
T382	M377	N101
A383	T382	E113
I384	A383	L136
D392	I384	S140
M398	D392	D160
F405	M398	L167
V405	F405	E169
H406	V405	F169
H407	H406	S170
	H407	I171
		V177
		S178
		T179
		A180
		V181
		H182
		E182

T234	M1
M235	C12
S236	F20
R243	I24
Q247	L30
I249	D81
A250	P32
D251	R48
K254	I49
L255	Y53
A256	L67
V257	V68
I258	D69
M259	V8
V260	T74
P261	V78
F262	I86
G263	F87
K264	F94
L285	G106
H266	E110
P270	G111
C305	H139
T314	M149
V315	G150
A316	T151
M325	E159
K326	E160
E327	F169
L333	S170
M325	H192
K326	Q193
E327	T201
L333	L204
M325	D205
K326	L209
E327	H229
L333	L230
M325	V231
K326	Q436

4 Experimental information

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of segments used	17069	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TITAN	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	27500	Depositor
Image detector	Not provided	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, TA1

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	0.40	0/3419	0.66	1/4639 (0.0%)
2	B	0.41	0/3427	0.66	0/4641
All	All	0.41	0/6846	0.66	1/9280 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	178	SER	N-CA-CB	6.32	119.98	110.50

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	3343	0	3262	119	0
2	B	3352	0	3235	127	0
3	A	32	0	12	0	0
4	A	1	0	0	0	0
5	B	28	0	12	0	0
6	B	62	0	51	9	0
All	All	6818	0	6572	151	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 11.

The worst 5 of 151 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:407:TRP:CG	2:B:257:VAL:HG22	1.18	1.66
1:A:210:TYR:CG	2:B:326:LYS:HD3	1.40	1.52
1:A:177:VAL:HG22	2:B:333:LEU:CD1	1.40	1.47
1:A:398:MET:SD	2:B:348:PRO:HD2	1.52	1.46
1:A:407:TRP:CD2	2:B:257:VAL:HG22	1.56	1.38

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	422/437 (97%)	389 (92%)	31 (7%)	2 (0%)	34	78
2	B	424/426 (100%)	412 (97%)	11 (3%)	1 (0%)	52	88
All	All	846/863 (98%)	801 (95%)	42 (5%)	3 (0%)	43	81

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	369	ARG
1	A	113	GLU
1	A	240	ALA

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	359/368 (98%)	350 (98%)	9 (2%)	55	84
2	B	367/367 (100%)	351 (96%)	16 (4%)	35	73
All	All	726/735 (99%)	701 (97%)	25 (3%)	48	79

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	139	HIS
2	B	159	GLU
2	B	356	CYS
2	B	149	MET
2	B	160	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	249	ASN
2	B	350	ASN
2	B	309	HIS
1	A	256	GLN
2	B	349	ASN

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$
3	GTP	A	501	4	26,34,34	1.00	2 (7%)	29,54,54	1.64	5 (17%)
5	GDP	B	501	-	24,30,30	1.22	3 (12%)	26,47,47	1.67	6 (23%)
6	TA1	B	502	-	68,68,68	0.51	0	102,105,105	0.66	1 (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
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
6	TA1	B	502	-	-	0/41/127/127	0/5/7/7

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	501	GDP	O4'-C1'	2.16	1.44	1.41
3	A	501	GTP	C5-C4	2.77	1.46	1.40
5	B	501	GDP	C5-C4	3.03	1.47	1.40
3	A	501	GTP	C6-C5	3.18	1.47	1.41
5	B	501	GDP	C6-C5	3.80	1.49	1.41

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	501	GTP	C5-C6-N1	-3.89	118.44	123.52
5	B	501	GDP	C5-C6-N1	-3.50	118.95	123.52
5	B	501	GDP	N3-C2-N1	-3.48	122.82	127.56
3	A	501	GTP	N3-C2-N1	-3.38	122.95	127.56

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	501	GTP	C6-C5-C4	-2.97	117.46	120.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	502	TA1	9	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.