wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:48 PM BST

PDB ID	:	3J7I
EMDB ID:	:	EMD-2697
Title	:	Structure of alpha- and beta- tubulin in GMPCPP-microtubules
Authors	:	Yajima, H.; Ogura, T.; Nitta, R.; Okada, Y.; Sato, C.; Hirokawa, N.
Deposited on	:	2014-07-01
Resolution	:	8.90 Å(reported)
Based on PDB ID	:	1JFF

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

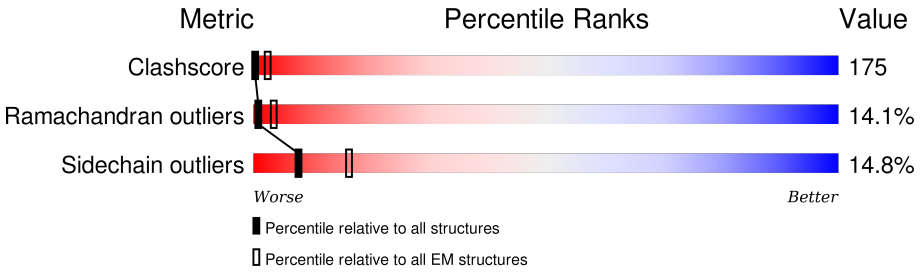
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

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 8.90 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	12% 57% 20% • 9%
2	B	445	19% 55% 15% • 8%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	GTP	A	502	-	-	X	-
4	GTP	B	502	-	-	X	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6515 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-1A chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	409	Total	C	N	O	S	0	0
			3210	2034	548	608	20		

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	265	GLY	ALA	SEE REMARK 999	UNP P02550

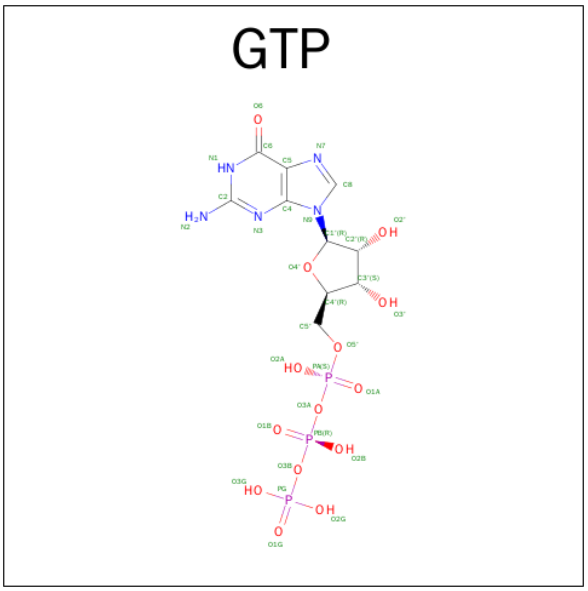
- Molecule 2 is a protein called Tubulin beta chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	411	Total	C	N	O	S	0	0
			3239	2037	558	620	24		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
3	B	1	Total	Mg	0
			1	1	
3	A	1	Total	Mg	0
			1	1	

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



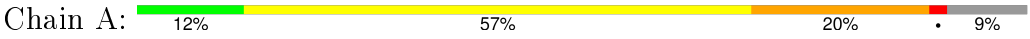
Mol	Chain	Residues	Atoms					AltConf
4	A	1	Total	C	N	O	P	0
			32	10	5	14	3	
4	B	1	Total	C	N	O	P	0
			32	10	5	14	3	

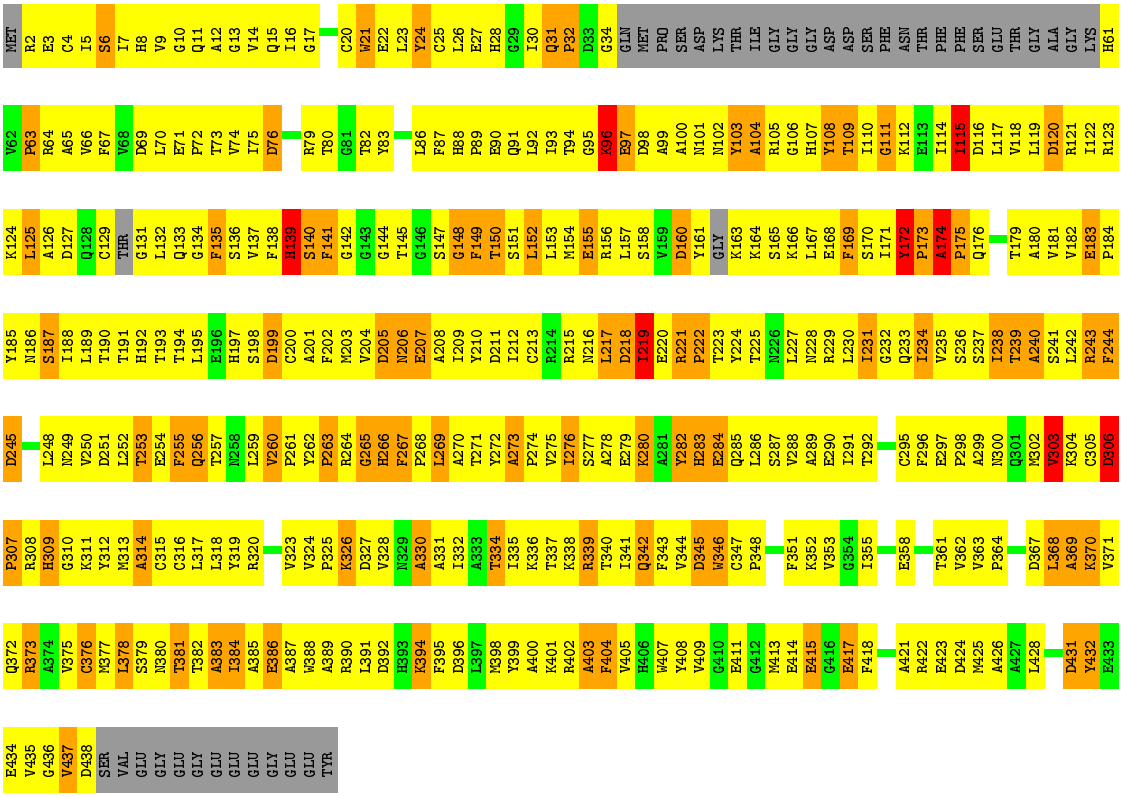
3 Residue-property plots (i)

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-1A chain

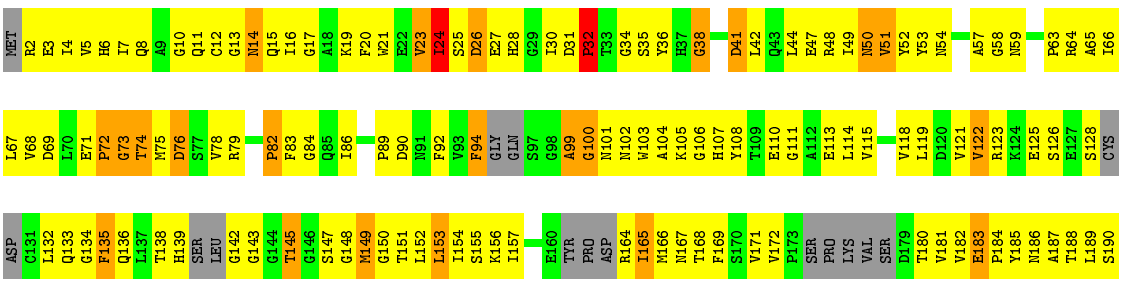
Chain A:

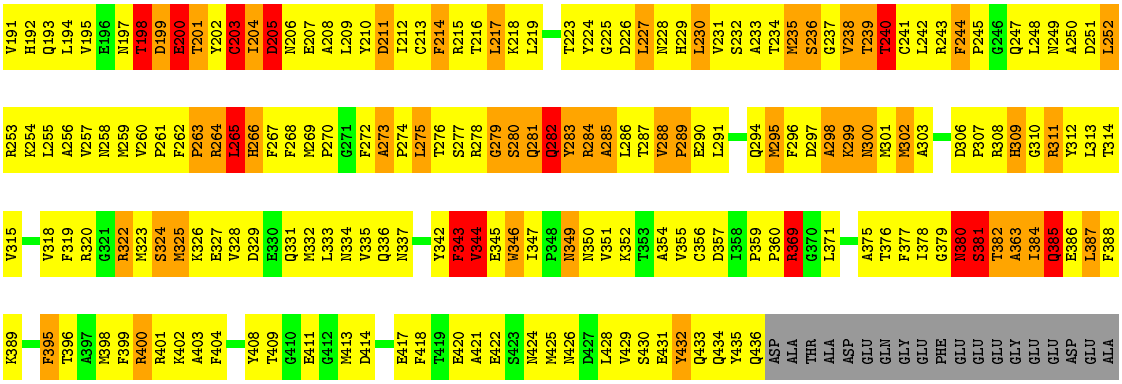




- Molecule 2: Tubulin beta chain

Chain B:




V191 H192 Q193 L194 V195 E196 N197 T198 D199 E200 T201 Y202 C203 I204 D205 N206 E207 A208 L209 Y210 D211 I212 C213 F214 R215 T216 L217 K218 L219 T223 Y224 G225 D226 L227 N228 H229 L230 V231 S232 A233 T234 M235 S236 G237 V238 T239 T240 C241 L242 R243 F244 P245 G246 Q247 L248 N249 A250 D251 L252
R253 K254 L255 A256 V257 N258 M259 V260 P261 F262 P263 R264 L265 H266 F267 F268 M269 P270 G271 F272 A273 P274 L275 T276 S277 R278 G279 S280 Q281 Q282 Y283 R284 A285 L286 T287 V288 P289 E290 L291 Q294 M295 F296 D297 A298 K299 N300 M301 M302 A303 D306 P307 R308 H309 G310 R311 Y312 L313 T314
V315 V318 F319 R320 G321 R322 M323 S324 M325 K326 E327 V328 D329 E330 Q331 M332 L333 N334 V335 Q336 N337 Y342 F343 V344 E345 W346 I347 P348 N349 N350 V351 K352 T353 A354 V355 G356 D357 I358 P359 P360 R369 G370 L371 A375 T376 F377 I378 G379 N380 S381 T382 A363 I384 Q385 E386 L387 F388
K389 F395 T396 A397 M398 F399 R400 R401 K402 A403 F404 Y408 T409 G410 E411 G412 M413 D414 E417 F418 T419 E420 A421 E422 S423 N424 M425 N426 D427 L428 V429 S430 E431 Y432 Q433 Q434 Y435 Q436 ASP ALA THR 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	Not provided	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	320000	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	Each Filament	Depositor
Microscope	JEM-2010F	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	10	Depositor
Minimum defocus (nm)	1300	Depositor
Maximum defocus (nm)	2800	Depositor
Magnification	40000	Depositor
Image detector	Kodak SO163 Film	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	0.55	2/3281 (0.1%)	0.78	10/4453 (0.2%)
2	B	0.57	3/3306 (0.1%)	0.87	15/4469 (0.3%)
All	All	0.56	5/6587 (0.1%)	0.83	25/8922 (0.3%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
2	B	0	1
All	All	0	2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	205	ASP	CB-CG	-9.46	1.31	1.51
2	B	205	ASP	CA-CB	-6.95	1.38	1.53
1	A	173	PRO	N-CD	5.43	1.55	1.47
1	A	307	PRO	N-CD	5.16	1.55	1.47
2	B	205	ASP	CA-C	5.08	1.66	1.52

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	205	ASP	CB-CG-OD1	-9.38	109.85	118.30
2	B	205	ASP	CA-CB-CG	-9.22	93.12	113.40
1	A	383	ALA	CA-C-O	-8.86	101.49	120.10
2	B	203	CYS	C-N-CA	8.78	143.65	121.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	205	ASP	C-N-CA	-8.20	101.19	121.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	202	PHE	Mainchain
2	B	380	ASN	Mainchain

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	3210	0	3123	1163	0
2	B	3239	0	3118	1137	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	32	0	12	61	0
4	B	32	0	12	27	0
All	All	6515	0	6265	2233	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 175.

The worst 5 of 2233 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:154:MET:HE2	1:A:197:HIS:CE1	1.22	1.68
2:B:185:TYR:CE1	2:B:399:PHE:HD1	1.02	1.63
1:A:189:LEU:HD21	1:A:418:PHE:CD2	1.12	1.63
2:B:413:MET:HE3	2:B:418:PHE:CE1	1.25	1.62
2:B:169:PHE:CZ	2:B:235:MET:CB	1.83	1.62

There are no symmetry-related clashes.

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 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	401/451 (89%)	257 (64%)	83 (21%)	61 (15%)	0 5
2	B	399/445 (90%)	262 (66%)	85 (21%)	52 (13%)	0 7
All	All	800/896 (89%)	519 (65%)	168 (21%)	113 (14%)	1 6

5 of 113 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	96	LYS
1	A	108	TYR
1	A	109	THR
1	A	175	PRO
1	A	183	GLU

5.3.2 Protein sidechains

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	345/377 (92%)	293 (85%)	52 (15%)	3 23
2	B	353/381 (93%)	302 (86%)	51 (14%)	4 25
All	All	698/758 (92%)	595 (85%)	103 (15%)	8 24

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	394	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	68	VAL
2	B	344	VAL
1	A	404	PHE
2	B	14	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	101	ASN
2	B	136	GLN
2	B	349	ASN
2	B	102	ASN
2	B	107	HIS

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 4 ligands modelled in this entry, 2 are 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	$\#\|Z\|>2$
4	GTP	A	502	3	26,34,34	1.36	2 (7%)	29,54,54	2.30	4 (13%)
4	GTP	B	502	3	26,34,34	1.35	1 (3%)	29,54,54	2.29	4 (13%)

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	502	3	-	0/18/38/38	0/3/3/3
4	GTP	B	502	3	-	0/18/38/38	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	502	GTP	O4'-C1'	2.01	1.44	1.41
4	B	502	GTP	C6-N1	4.81	1.41	1.33
4	A	502	GTP	C6-N1	4.82	1.41	1.33

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	502	GTP	C5-C6-N1	-7.81	113.31	123.52
4	B	502	GTP	C5-C6-N1	-7.80	113.33	123.52
4	A	502	GTP	N3-C2-N1	-3.59	122.67	127.56
4	B	502	GTP	N3-C2-N1	-3.49	122.81	127.56
4	B	502	GTP	C6-C5-C4	-2.14	118.41	120.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 88 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	502	GTP	61	0
4	B	502	GTP	27	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.