



## wwPDB EM Map/Model Validation Report ⓘ

Apr 10, 2016 – 02:03 PM BST

PDB ID : 4ATU  
EMDB ID: : EMD-2095  
Title : Human doublecortin N-DC repeat plus linker, and tubulin (2XRP) docked into an 8Å cryo-EM map of doublecortin-stabilised microtubules reconstructed in absence of kinesin  
Authors : Liu, J.S.; Schubert, C.R.; Fu, X.; Fourniol, F.J.; Jaiswal, J.K.; Houdusse, A.; Stultz, C.M.; Moores, C.A.; Walsh, C.A.  
Deposited on : 2012-05-09  
Resolution : 8.30 Å(reported)  
Based on PDB ID : 2XRP

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](mailto: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

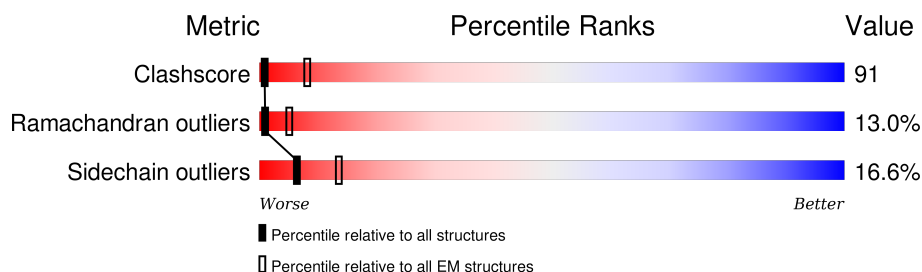
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 8.30 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  $\geq 3$ , 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	445	21% 55% 18% . .
1	C	445	19% 57% 18% . .
1	E	445	21% 55% 18% . .
1	G	445	19% 56% 18% . .
2	B	452	16% 59% 19% . 5%
2	D	452	17% 58% 18% . 5%
2	F	452	17% 58% 18% . 5%
2	H	452	17% 58% 17% . 5%
3	I	373	10% 13% . . 71%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 27840 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 BETA-2B CHAIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	A	426	Total	C	N	O	S	0	0
			3351	2105	575	646	25		
1	C	426	Total	C	N	O	S	0	0
			3351	2105	575	646	25		
1	E	426	Total	C	N	O	S	0	0
			3351	2105	575	646	25		
1	G	426	Total	C	N	O	S	0	0
			3351	2105	575	646	25		

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	57	ALA	THR	CONFLICT	UNP Q6B856
A	172	VAL	MET	CONFLICT	UNP Q6B856
A	298	ALA	SER	CONFLICT	UNP Q6B856
A	318	VAL	ILE	CONFLICT	UNP Q6B856
C	57	ALA	THR	CONFLICT	UNP Q6B856
C	172	VAL	MET	CONFLICT	UNP Q6B856
C	298	ALA	SER	CONFLICT	UNP Q6B856
C	318	VAL	ILE	CONFLICT	UNP Q6B856
E	57	ALA	THR	CONFLICT	UNP Q6B856
E	172	VAL	MET	CONFLICT	UNP Q6B856
E	298	ALA	SER	CONFLICT	UNP Q6B856
E	318	VAL	ILE	CONFLICT	UNP Q6B856
G	57	ALA	THR	CONFLICT	UNP Q6B856
G	172	VAL	MET	CONFLICT	UNP Q6B856
G	298	ALA	SER	CONFLICT	UNP Q6B856
G	318	VAL	ILE	CONFLICT	UNP Q6B856

- Molecule 2 is a protein called TUBULIN ALPHA-1D CHAIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	429	Total	C	N	O	S	0	0
			3334	2114	569	630	21		
2	D	429	Total	C	N	O	S	0	0
			3334	2114	569	630	21		
2	F	429	Total	C	N	O	S	0	0
			3334	2114	569	630	21		
2	H	429	Total	C	N	O	S	0	0
			3334	2114	569	630	21		

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	450	GLU	ASP	CONFLICT	UNP Q2HJ86
B	7	ILE	VAL	CONFLICT	UNP Q2HJ86
B	114	ILE	LEU	CONFLICT	UNP Q2HJ86
B	136	SER	LEU	CONFLICT	UNP Q2HJ86
B	137	VAL	ILE	CONFLICT	UNP Q2HJ86
B	265	GLY	ILE	CONFLICT	UNP Q2HJ86
B	358	GLU	GLN	CONFLICT	UNP Q2HJ86
B	437	VAL	MET	CONFLICT	UNP Q2HJ86
D	450	GLU	ASP	CONFLICT	UNP Q2HJ86
D	7	ILE	VAL	CONFLICT	UNP Q2HJ86
D	114	ILE	LEU	CONFLICT	UNP Q2HJ86
D	136	SER	LEU	CONFLICT	UNP Q2HJ86
D	137	VAL	ILE	CONFLICT	UNP Q2HJ86
D	265	GLY	ILE	CONFLICT	UNP Q2HJ86
D	358	GLU	GLN	CONFLICT	UNP Q2HJ86
D	437	VAL	MET	CONFLICT	UNP Q2HJ86
F	450	GLU	ASP	CONFLICT	UNP Q2HJ86
F	7	ILE	VAL	CONFLICT	UNP Q2HJ86
F	114	ILE	LEU	CONFLICT	UNP Q2HJ86
F	136	SER	LEU	CONFLICT	UNP Q2HJ86
F	137	VAL	ILE	CONFLICT	UNP Q2HJ86
F	265	GLY	ILE	CONFLICT	UNP Q2HJ86
F	358	GLU	GLN	CONFLICT	UNP Q2HJ86
F	437	VAL	MET	CONFLICT	UNP Q2HJ86
H	450	GLU	ASP	CONFLICT	UNP Q2HJ86
H	7	ILE	VAL	CONFLICT	UNP Q2HJ86
H	114	ILE	LEU	CONFLICT	UNP Q2HJ86
H	136	SER	LEU	CONFLICT	UNP Q2HJ86
H	137	VAL	ILE	CONFLICT	UNP Q2HJ86
H	265	GLY	ILE	CONFLICT	UNP Q2HJ86
H	358	GLU	GLN	CONFLICT	UNP Q2HJ86
H	437	VAL	MET	CONFLICT	UNP Q2HJ86

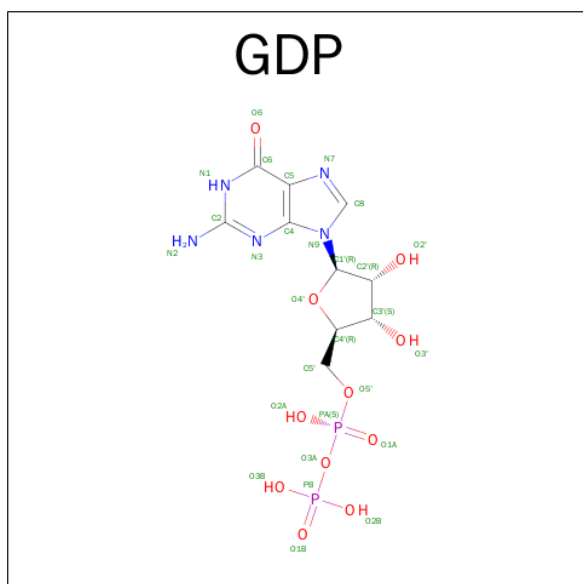
- Molecule 3 is a protein called NEURONAL MIGRATION PROTEIN DOUBLECORTIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	I	107	Total	C	N	O	S	0	1
			860	541	151	165	3		

There are 14 discrepancies between the modelled and reference sequences:

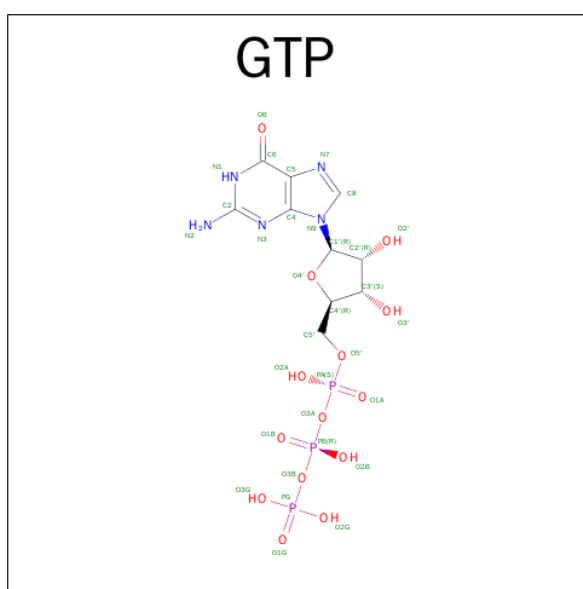
Chain	Residue	Modelled	Actual	Comment	Reference
I	-12	GLY	-	EXPRESSION TAG	UNP O43602
I	-11	ALA	-	EXPRESSION TAG	UNP O43602
I	-10	MET	-	EXPRESSION TAG	UNP O43602
I	-9	ASP	-	EXPRESSION TAG	UNP O43602
I	-8	TYR	-	EXPRESSION TAG	UNP O43602
I	-7	LYS	-	EXPRESSION TAG	UNP O43602
I	-6	ASP	-	EXPRESSION TAG	UNP O43602
I	-5	ASP	-	EXPRESSION TAG	UNP O43602
I	-4	ASP	-	EXPRESSION TAG	UNP O43602
I	-3	ASP	-	EXPRESSION TAG	UNP O43602
I	-2	LYS	-	EXPRESSION TAG	UNP O43602
I	-1	ILE	-	EXPRESSION TAG	UNP O43602
I	0	CYS	-	EXPRESSION TAG	UNP O43602
I	1	ARG	-	EXPRESSION TAG	UNP O43602

- Molecule 4 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula:  $C_{10}H_{15}N_5O_{11}P_2$ ).



Mol	Chain	Residues	Atoms					AltConf
4	A	1	Total	C	N	O	P	0
			28	10	5	11	2	
4	C	1	Total	C	N	O	P	0
			28	10	5	11	2	
4	E	1	Total	C	N	O	P	0
			28	10	5	11	2	
4	G	1	Total	C	N	O	P	0
			28	10	5	11	2	

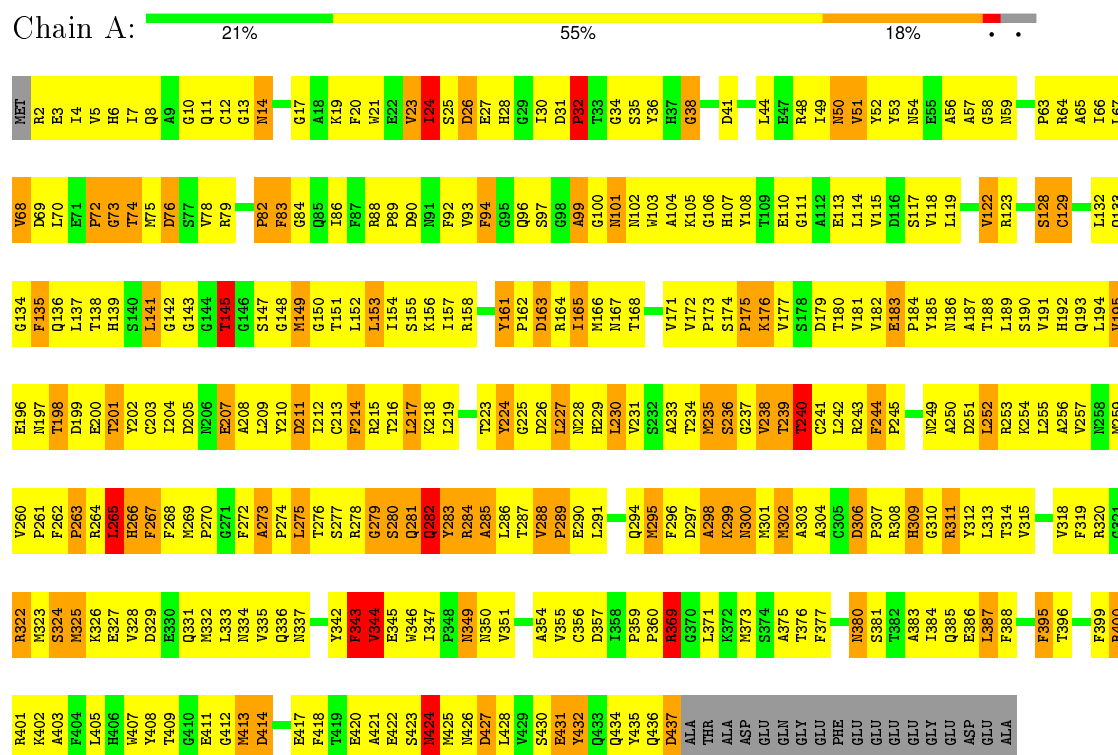
- Molecule 5 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:  $C_{10}H_{16}N_5O_{14}P_3$ ).



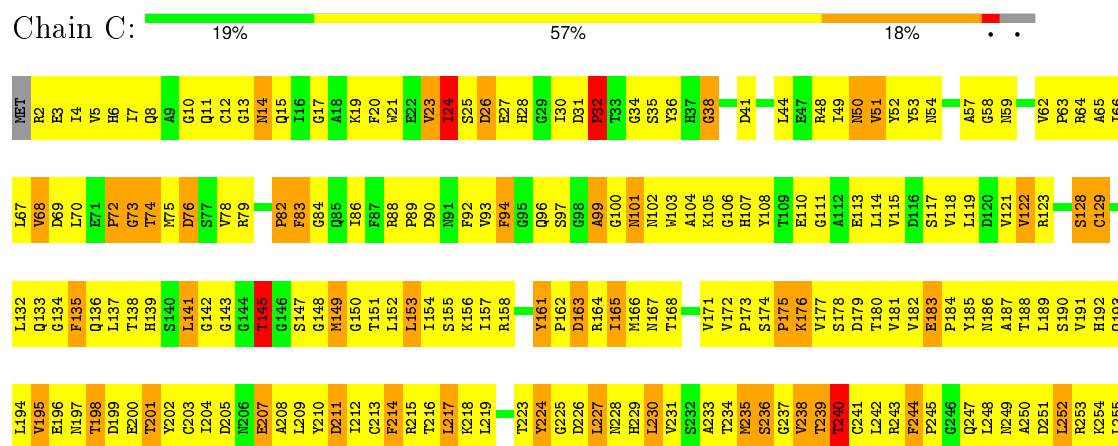
### 3 Residue-property plots

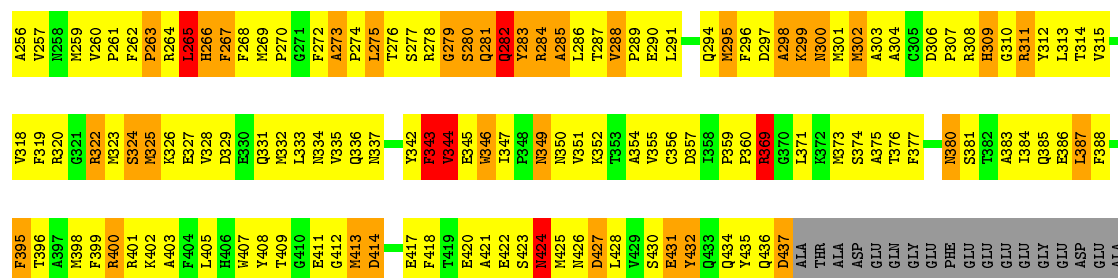
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric 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 BETA-2B CHAIN



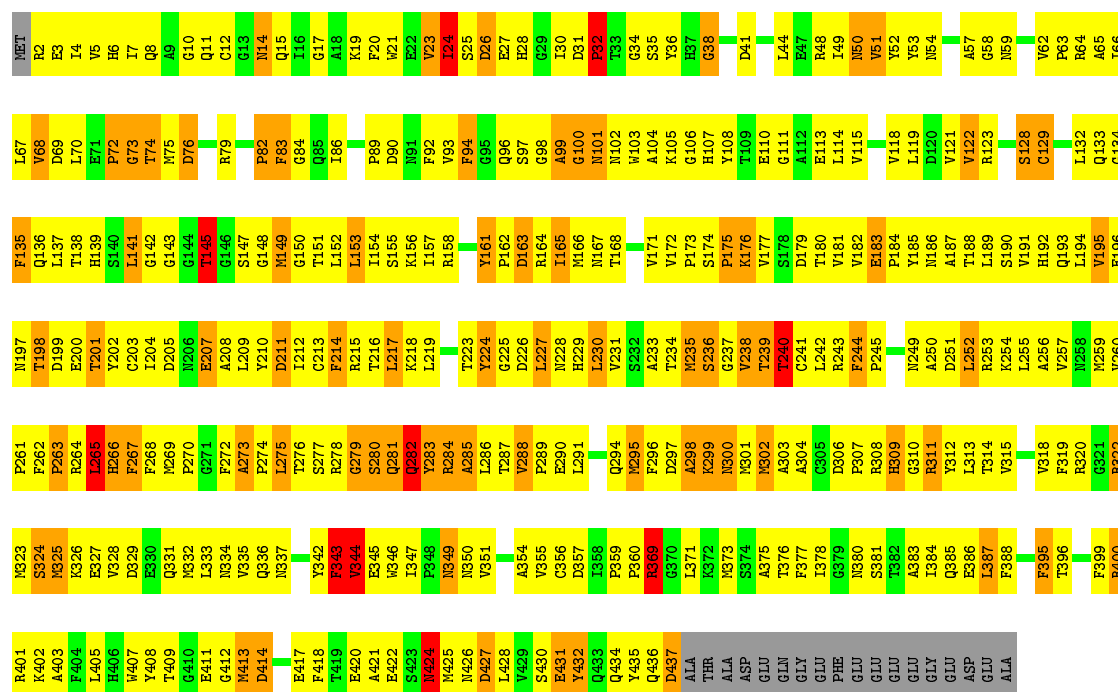
#### • Molecule 1: TUBULIN BETA-2B CHAIN





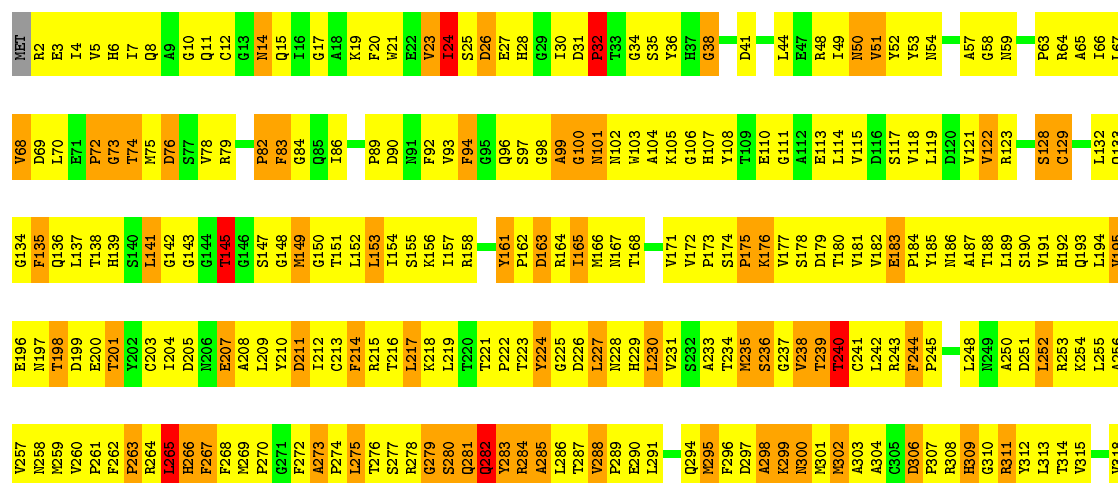
### • Molecule 1: TUBULIN BETA-2B CHAIN

Chain E: 21% 55% 18%

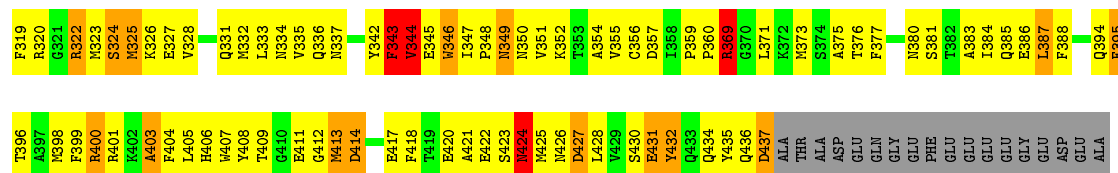


### • Molecule 1: TUBULIN BETA-2B CHAIN

Chain G: 19% 56% 18%

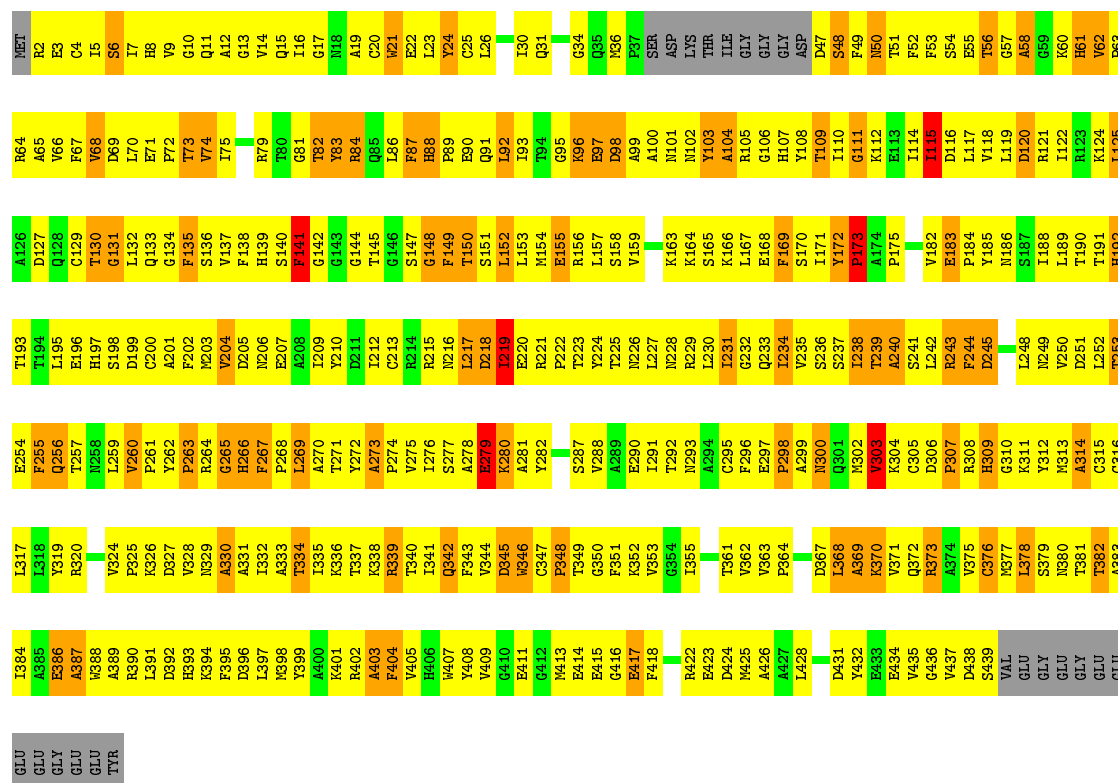






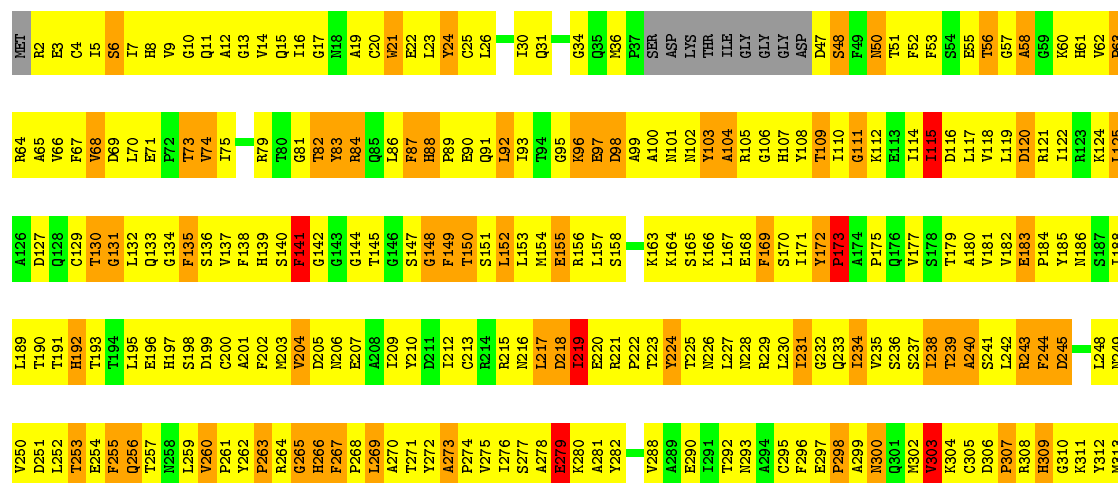
## • Molecule 2: TUBULIN ALPHA-1D CHAIN

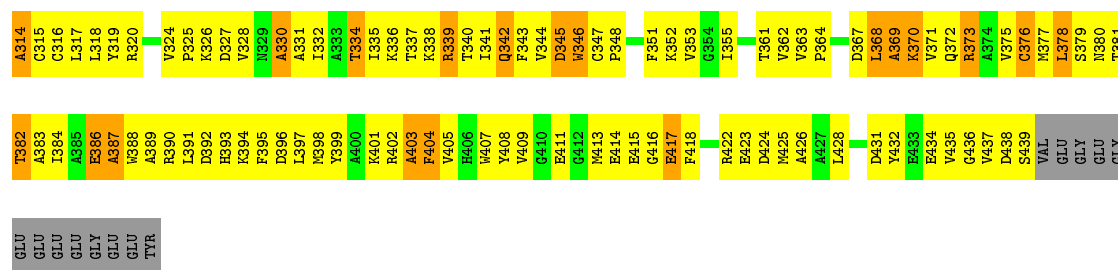
Chain B: 16% 59% 19% 5%



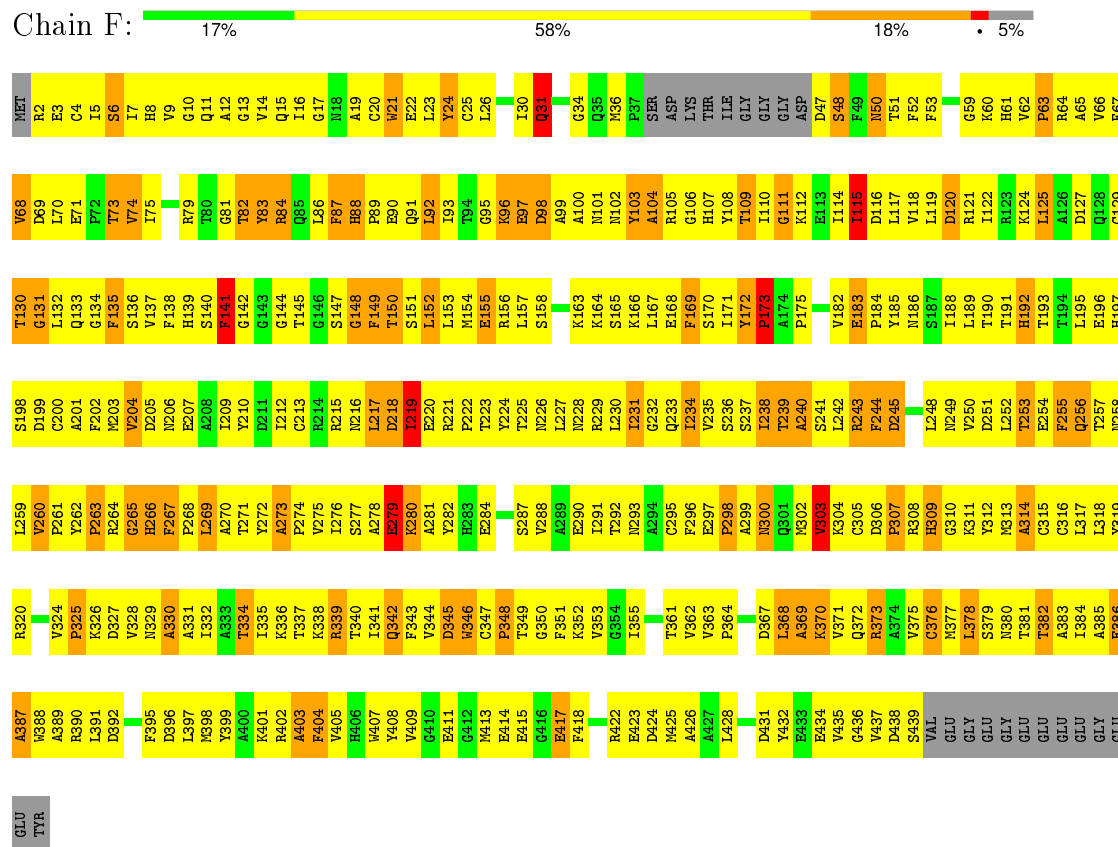
## • Molecule 2: TUBULIN ALPHA-1D CHAIN

Chain D: 17% 58% 18% 5%

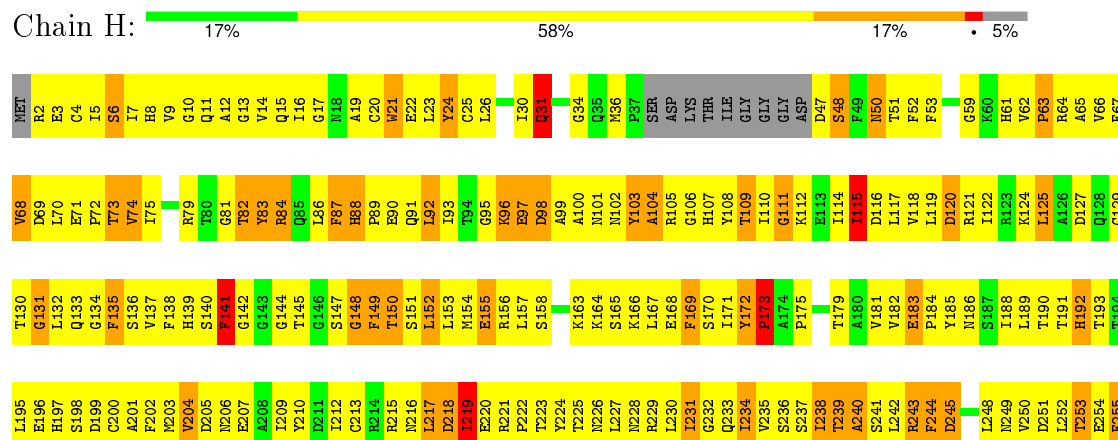




## • Molecule 2: TUBULIN ALPHA-1D CHAIN



## • Molecule 2: TUBULIN ALPHA-1D CHAIN





## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	DONE WITH FREALIGN, Not provided	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	17	Depositor
Minimum defocus (nm)	900	Depositor
Maximum defocus (nm)	3600	Depositor
Magnification	50000	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: GDP, GTP

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	0.51	0/3426	0.76	2/4642 (0.0%)
1	C	0.51	0/3426	0.76	2/4642 (0.0%)
1	E	0.51	0/3426	0.76	2/4642 (0.0%)
1	G	0.51	0/3426	0.76	2/4642 (0.0%)
2	B	0.82	3/3410 (0.1%)	0.77	3/4629 (0.1%)
2	D	0.82	3/3410 (0.1%)	0.78	3/4629 (0.1%)
2	F	0.82	3/3410 (0.1%)	0.78	3/4629 (0.1%)
2	H	0.82	3/3410 (0.1%)	0.78	3/4629 (0.1%)
3	I	0.82	0/876	1.21	4/1178 (0.3%)
All	All	0.69	12/28220 (0.0%)	0.78	24/38262 (0.1%)

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
3	I	0	8

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	92	LEU	C-N	27.98	1.98	1.34
2	F	92	LEU	C-N	27.96	1.98	1.34
2	H	92	LEU	C-N	27.95	1.98	1.34
2	B	92	LEU	C-N	27.93	1.98	1.34
2	H	298	PRO	C-N	17.32	1.73	1.34

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	I	102	ARG	NE-CZ-NH1	7.97	124.28	120.30
2	D	68	VAL	O-C-N	-7.15	111.26	122.70
2	F	68	VAL	O-C-N	-7.11	111.33	122.70
2	B	68	VAL	O-C-N	-7.10	111.34	122.70
2	H	68	VAL	O-C-N	-7.08	111.37	122.70

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	I	105	TYR	Sidechain
3	I	123	GLU	Sidechain
3	I	129	SER	Peptide
3	I	81	ASP	Sidechain
3	I	93	ASP	Sidechain

## 5.2 Too-close contacts

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	3351	0	3229	541	0
1	C	3351	0	3229	617	0
1	E	3351	0	3229	543	0
1	G	3351	0	3229	748	0
2	B	3334	0	3223	687	0
2	D	3334	0	3223	665	0
2	F	3334	0	3223	728	0
2	H	3334	0	3223	714	0
3	I	860	0	841	98	0
4	A	28	0	12	1	0
4	C	28	0	12	1	0
4	E	28	0	12	1	0
4	G	28	0	12	4	0
5	B	32	0	12	5	0
5	D	32	0	12	4	0
5	F	32	0	12	4	0
5	H	32	0	12	5	0
All	All	27840	0	26745	4989	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 91.

The worst 5 of 4989 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:G:257:VAL:HG21	2:H:407:TRP:CG	1.28	1.64
2:B:296:PHE:CE2	2:B:341:ILE:HD11	1.32	1.62
2:H:296:PHE:CE2	2:H:341:ILE:HD11	1.32	1.62
2:F:296:PHE:CE2	2:F:341:ILE:HD11	1.32	1.62
2:D:296:PHE:CE2	2:D:341:ILE:HD11	1.32	1.62

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	424/445 (95%)	274 (65%)	94 (22%)	56 (13%)	0	7
1	C	424/445 (95%)	273 (64%)	95 (22%)	56 (13%)	0	7
1	E	424/445 (95%)	273 (64%)	95 (22%)	56 (13%)	0	7
1	G	424/445 (95%)	273 (64%)	95 (22%)	56 (13%)	0	7
2	B	423/452 (94%)	279 (66%)	87 (21%)	57 (14%)	0	7
2	D	423/452 (94%)	281 (66%)	85 (20%)	57 (14%)	0	7
2	F	423/452 (94%)	279 (66%)	87 (21%)	57 (14%)	0	7
2	H	423/452 (94%)	278 (66%)	88 (21%)	57 (14%)	0	7
3	I	105/373 (28%)	88 (84%)	15 (14%)	2 (2%)	10	52
All	All	3493/3961 (88%)	2298 (66%)	741 (21%)	454 (13%)	1	7

5 of 454 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	23	VAL
1	A	24	ILE
1	A	32	PRO
1	A	50	ASN
1	A	82	PRO

### 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	367/381 (96%)	307 (84%)	60 (16%)	3	20
1	C	367/381 (96%)	308 (84%)	59 (16%)	3	20
1	E	367/381 (96%)	308 (84%)	59 (16%)	3	20
1	G	367/381 (96%)	308 (84%)	59 (16%)	3	20
2	B	354/378 (94%)	295 (83%)	59 (17%)	3	19
2	D	354/378 (94%)	297 (84%)	57 (16%)	3	20
2	F	354/378 (94%)	296 (84%)	58 (16%)	3	19
2	H	354/378 (94%)	298 (84%)	56 (16%)	3	21
3	I	95/329 (29%)	66 (70%)	29 (30%)	0	3
All	All	2979/3365 (88%)	2483 (83%)	496 (17%)	6	19

5 of 496 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	345	ASP
1	E	309	HIS
2	H	368	LEU
2	D	404	PHE
1	E	153	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 123 such sidechains are listed below:



Mol	Chain	Res	Type
2	D	216	ASN
1	E	331	GLN
2	H	139	HIS
2	D	256	GLN
1	E	101	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 ⓘ

8 ligands are modelled in this entry.

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	GDP	A	600	-	24,30,30	2.76	8 (33%)	26,47,47	3.30	10 (38%)
5	GTP	B	500	-	26,34,34	1.56	4 (15%)	29,54,54	2.42	6 (20%)
4	GDP	C	600	-	24,30,30	2.76	8 (33%)	26,47,47	3.30	10 (38%)
5	GTP	D	500	-	26,34,34	1.36	1 (3%)	29,54,54	2.32	5 (17%)
4	GDP	E	600	-	24,30,30	2.75	8 (33%)	26,47,47	3.30	10 (38%)
5	GTP	F	500	-	26,34,34	1.37	2 (7%)	29,54,54	2.31	5 (17%)
4	GDP	G	600	-	24,30,30	2.76	8 (33%)	26,47,47	3.30	10 (38%)
5	GTP	H	500	-	26,34,34	1.36	2 (7%)	29,54,54	2.35	5 (17%)

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	GDP	A	600	-	-	0/12/32/32	0/3/3/3
5	GTP	B	500	-	-	0/18/38/38	0/3/3/3
4	GDP	C	600	-	-	0/12/32/32	0/3/3/3
5	GTP	D	500	-	-	0/18/38/38	0/3/3/3
4	GDP	E	600	-	-	0/12/32/32	0/3/3/3
5	GTP	F	500	-	-	0/18/38/38	0/3/3/3
4	GDP	G	600	-	-	0/12/32/32	0/3/3/3
5	GTP	H	500	-	-	0/18/38/38	0/3/3/3

The worst 5 of 41 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	G	600	GDP	PB-O2B	-4.24	1.40	1.54
4	A	600	GDP	PB-O2B	-4.24	1.40	1.54
4	C	600	GDP	PB-O2B	-4.22	1.40	1.54
4	E	600	GDP	PB-O2B	-4.20	1.40	1.54
5	H	500	GTP	O4'-C1'	2.01	1.44	1.41

The worst 5 of 61 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	600	GDP	C6-C5-C4	-9.93	109.51	120.86
4	A	600	GDP	C6-C5-C4	-9.90	109.55	120.86
4	E	600	GDP	C6-C5-C4	-9.86	109.59	120.86
4	C	600	GDP	C6-C5-C4	-9.84	109.61	120.86
5	H	500	GTP	C5-C6-N1	-7.89	113.20	123.52

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 25 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	600	GDP	1	0
5	B	500	GTP	5	0
4	C	600	GDP	1	0
5	D	500	GTP	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	E	600	GDP	1	0
5	F	500	GTP	4	0
4	G	600	GDP	4	0
5	H	500	GTP	5	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.