



## wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Feb 21, 2018 – 05:05 pm GMT

PDB ID : 2XRP  
EMDB ID: : EMD-1788  
Title : Human Doublecortin N-DC Repeat (1MJD) and Mammalian Tubulin (1JFF and 3HKE) Docked into the 8-Angstrom Cryo-EM Map of Doublecortin- Stabilised Microtubules  
Authors : Fourniol, F.J.; Sindelar, C.V.; Amigues, B.; Clare, D.K.; Thomas, G.; Perderiset, M.; Francis, F.; Houdusse, A.; Moores, C.A.  
Deposited on : 2010-09-18  
Resolution : 8.20 Å(reported)  
Based on PDB ID : 1MJD, 3HKE, 1JFF

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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

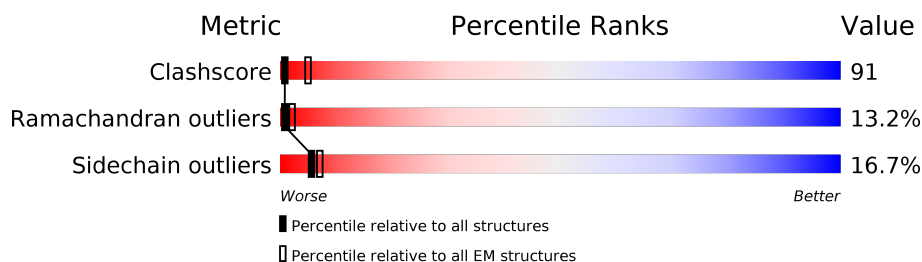
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 8.20 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	136279	1886
Ramachandran outliers	132675	1663
Sidechain outliers	132484	1531

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	20% 55% 18% . .
1	C	445	21% 55% 18% . .
1	E	445	22% 54% 18% . .
1	G	445	20% 56% 18% . .
2	B	452	16% 59% 19% . 5%
2	D	452	17% 59% 18% . 5%
2	F	452	17% 58% 18% . 5%
2	H	452	17% 58% 17% . 5%
3	I	95	32% 56% 13%

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 28352 atoms, of which 596 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 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	172	VAL	MET	conflict	UNP Q6B856
A	318	VAL	ILE	conflict	UNP Q6B856
C	172	VAL	MET	conflict	UNP Q6B856
C	318	VAL	ILE	conflict	UNP Q6B856
E	172	VAL	MET	conflict	UNP Q6B856
E	318	VAL	ILE	conflict	UNP Q6B856
G	172	VAL	MET	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 24 discrepancies between the modelled and reference sequences:

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

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

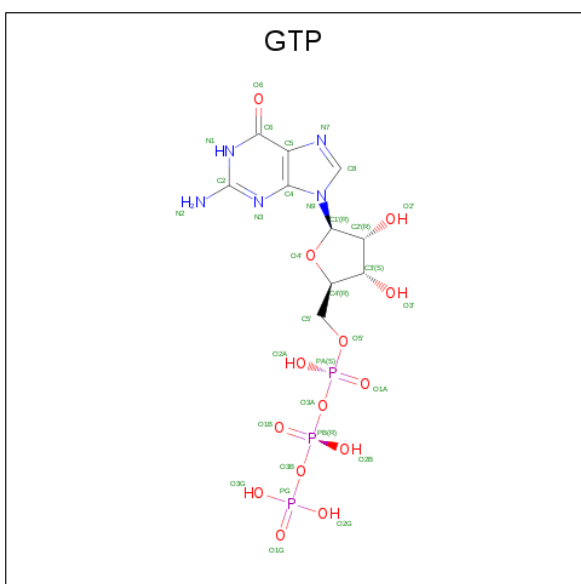
Mol	Chain	Residues	Atoms						AltConf	Trace
3	I	95	Total	C	H	N	O	S	0	0
			1372	490	596	134	150	2		

- 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 28	C 10	N 5	O 11	P 2	0
4	C	1	Total 28	C 10	N 5	O 11	P 2	0
4	E	1	Total 28	C 10	N 5	O 11	P 2	0
4	G	1	Total 28	C 10	N 5	O 11	P 2	0

- Molecule 5 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:  $\text{C}_{10}\text{H}_{16}\text{N}_5\text{O}_{14}\text{P}_3$ ).

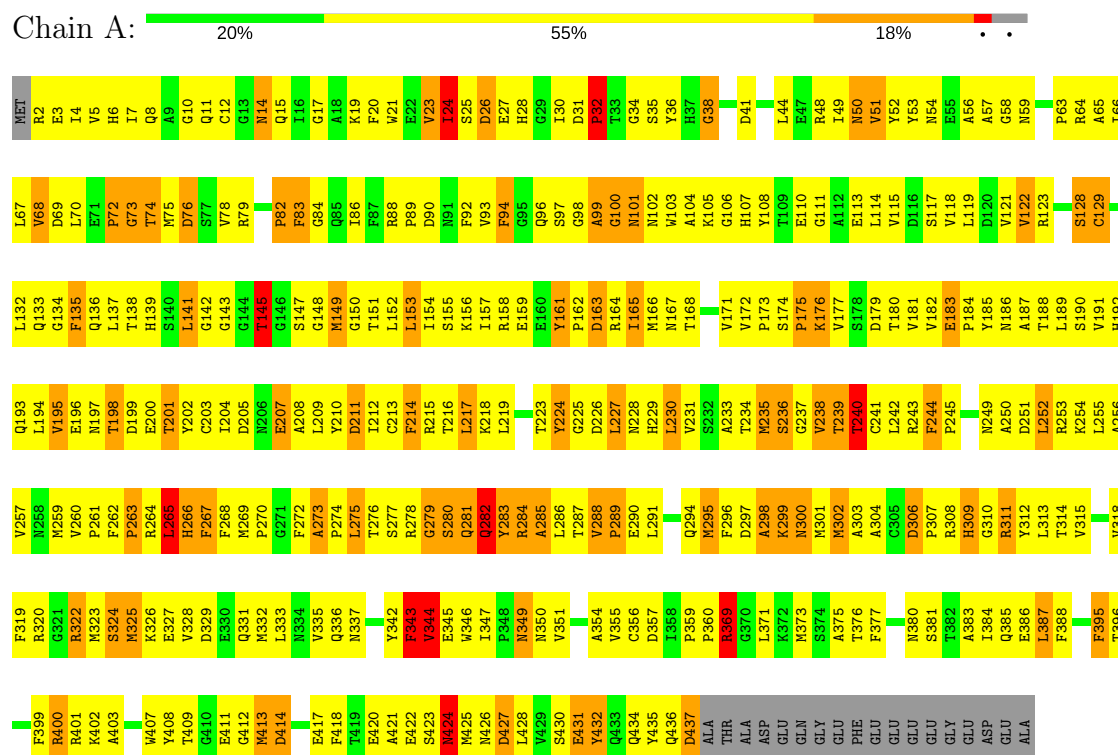


Mol	Chain	Residues	Atoms					AltConf
5	B	1	Total 32	C 10	N 5	O 14	P 3	0
5	D	1	Total 32	C 10	N 5	O 14	P 3	0
5	F	1	Total 32	C 10	N 5	O 14	P 3	0
5	H	1	Total 32	C 10	N 5	O 14	P 3	0

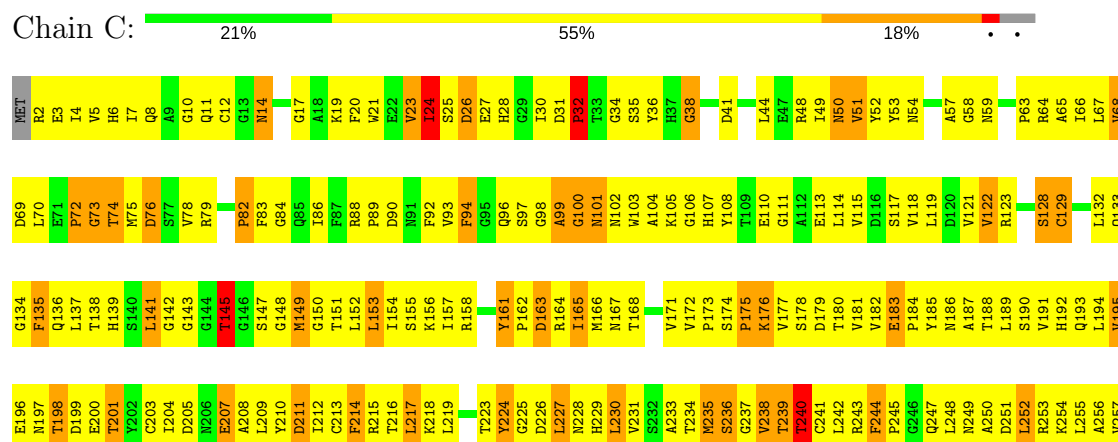
### 3 Residue-property plots

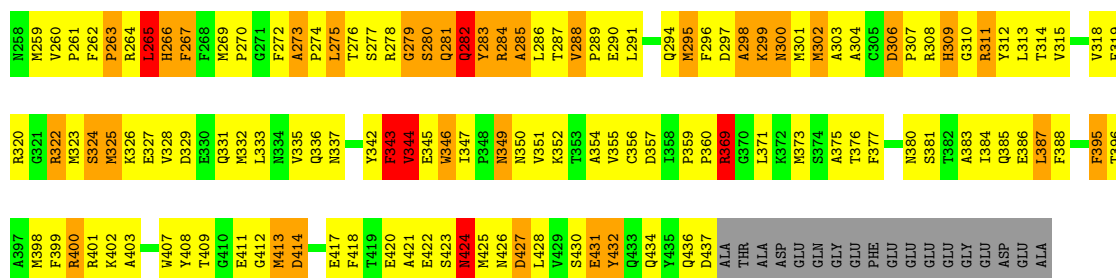
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according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



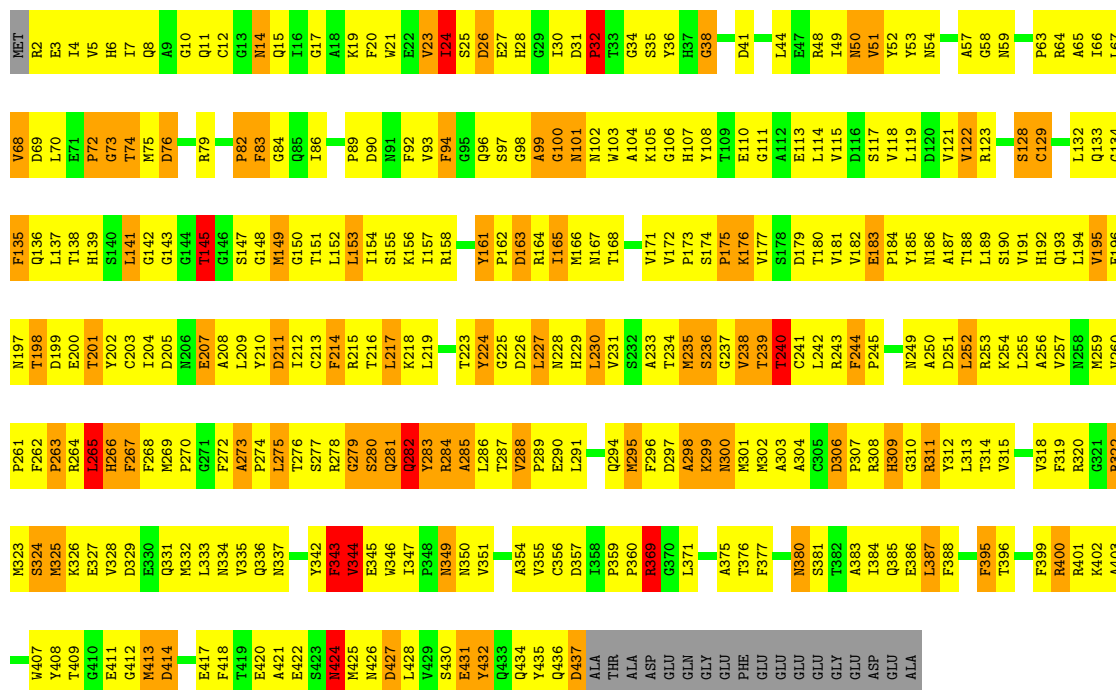
#### • Molecule 1: TUBULIN BETA-2B CHAIN





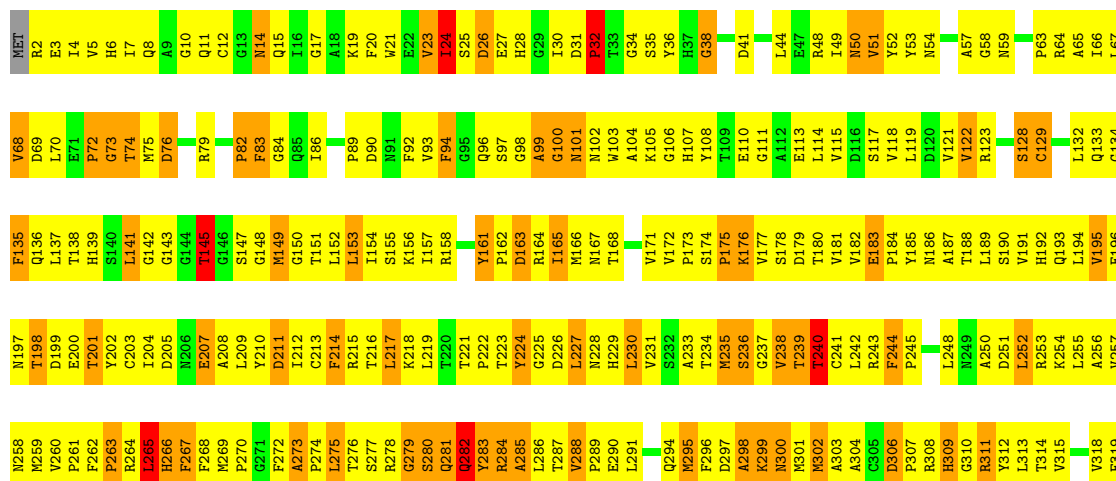
• Molecule 1: TUBULIN BETA-2B CHAIN

Chain E: 22% 54% 18%

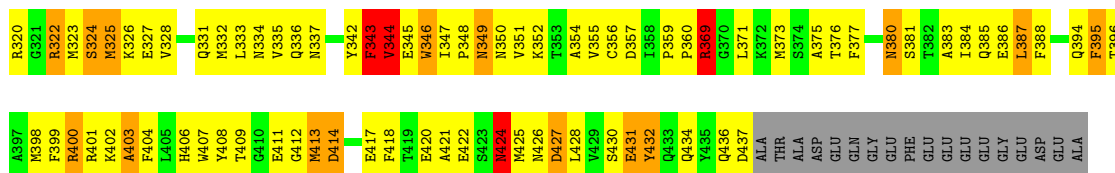


• Molecule 1: TUBULIN BETA-2B CHAIN

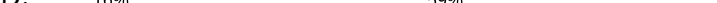
Chain G: 20% 56% 18%

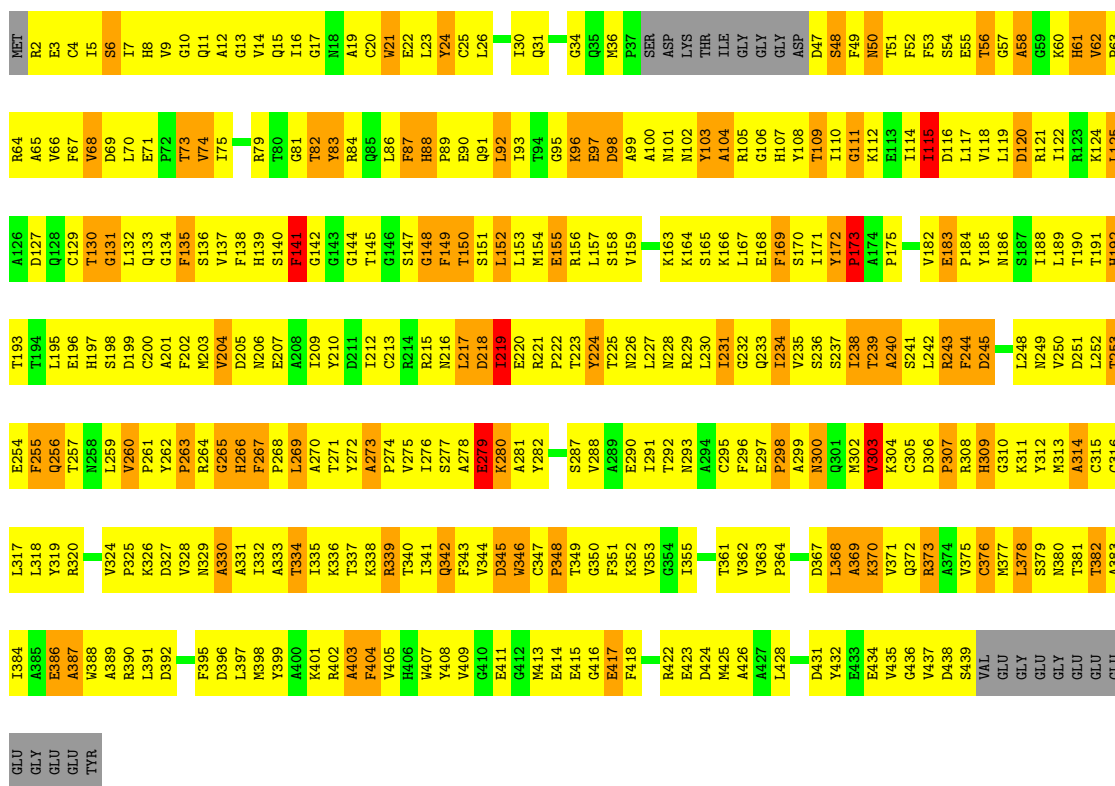




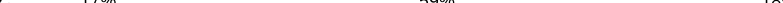


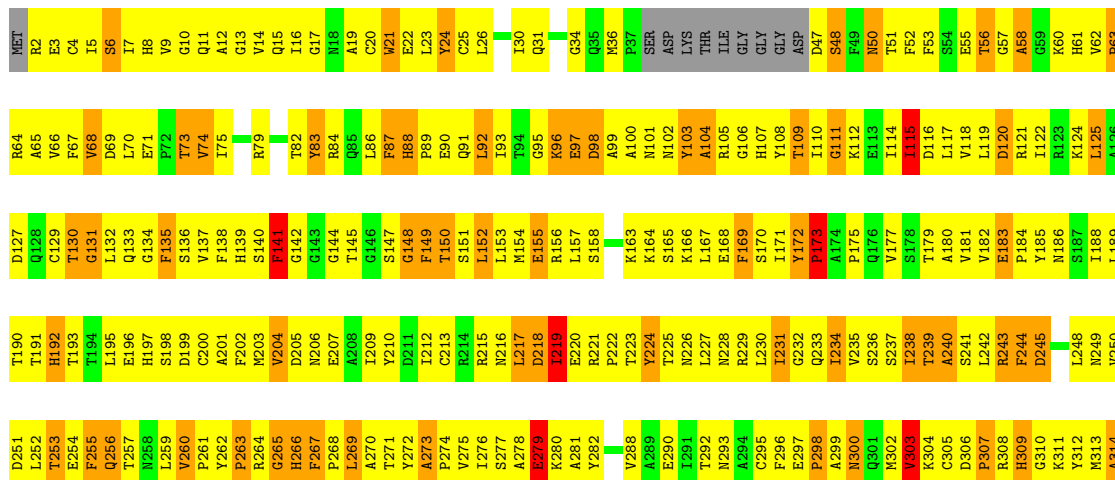
- Molecule 2: TUBULIN ALPHA-1D CHAIN

Chain B:  16% 59% 19% 5%

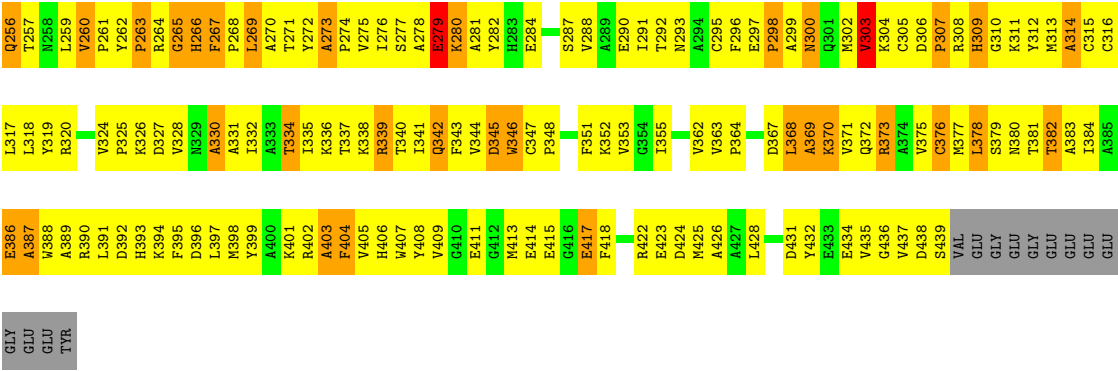


- Molecule 2: TUBULIN ALPHA-1D CHAIN

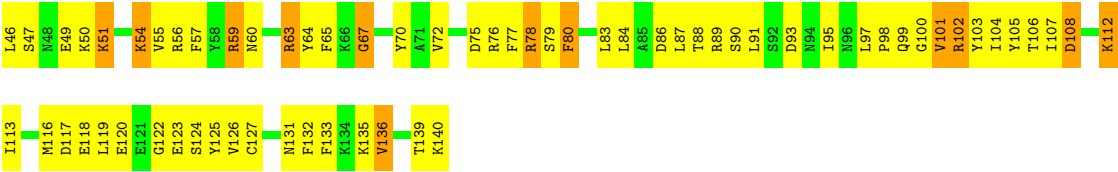
Chain D:  17% 59% 18% • 5%







● Molecule 3: NEURONAL MIGRATION PROTEIN DOUBLECORTIN



## 4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	HELICAL, twist=Not provided°, rise=Not provided Å, axial sym=Not provided	Depositor
Number of particles used	168000	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	DONE IN FREALIGN	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	15	Depositor
Minimum defocus (nm)	760	Depositor
Maximum defocus (nm)	2900	Depositor
Magnification	50000	Depositor
Image detector	KODAK SO-163 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	0.27	0/789	0.37	0/1055
All	All	0.67	12/28133 (0.0%)	0.76	20/38139 (0.1%)

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	27.99	1.98	1.34
2	B	92	LEU	C-N	27.96	1.98	1.34
2	H	92	LEU	C-N	27.95	1.98	1.34
2	F	92	LEU	C-N	27.95	1.98	1.34
2	F	298	PRO	C-N	17.32	1.73	1.34

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	68	VAL	O-C-N	-7.14	111.27	122.70
2	B	68	VAL	O-C-N	-7.12	111.31	122.70
2	F	68	VAL	O-C-N	-7.12	111.31	122.70
2	H	68	VAL	O-C-N	-7.10	111.34	122.70
2	F	298	PRO	O-C-N	-6.73	111.93	122.70

There are no chirality outliers.

There are no planarity outliers.

## 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	545	0
1	C	3351	0	3229	612	0
1	E	3351	0	3229	540	0
1	G	3351	0	3229	745	0
2	B	3334	0	3223	687	0
2	D	3334	0	3223	659	0
2	F	3334	0	3223	724	0
2	H	3334	0	3223	716	0
3	I	776	596	764	57	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	5	0
5	H	32	0	12	5	0
All	All	27756	596	26668	4935	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 4935 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	1.63
2:D:296:PHE:CE2	2:D:341:ILE:HD11	1.32	1.59
2:B:296:PHE:CE2	2:B:341:ILE:HD11	1.32	1.59
2:F:296:PHE:CE2	2:F:341:ILE:HD11	1.32	1.57
2:H:296:PHE:CE2	2:H:341:ILE:HD11	1.32	1.57

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	6
1	C	424/445 (95%)	274 (65%)	94 (22%)	56 (13%)	0	6
1	E	424/445 (95%)	274 (65%)	94 (22%)	56 (13%)	0	6
1	G	424/445 (95%)	274 (65%)	94 (22%)	56 (13%)	0	6
2	B	423/452 (94%)	279 (66%)	87 (21%)	57 (14%)	0	5
2	D	423/452 (94%)	281 (66%)	85 (20%)	57 (14%)	0	5
2	F	423/452 (94%)	279 (66%)	87 (21%)	57 (14%)	0	5
2	H	423/452 (94%)	278 (66%)	88 (21%)	57 (14%)	0	5
3	I	93/95 (98%)	72 (77%)	15 (16%)	6 (6%)	1	21
All	All	3481/3683 (94%)	2285 (66%)	738 (21%)	458 (13%)	1	6

5 of 458 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 [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	367/381 (96%)	307 (84%)	60 (16%)	2	15
1	C	367/381 (96%)	308 (84%)	59 (16%)	2	16
1	E	367/381 (96%)	308 (84%)	59 (16%)	2	16
1	G	367/381 (96%)	308 (84%)	59 (16%)	2	16
2	B	354/378 (94%)	295 (83%)	59 (17%)	2	15
2	D	354/378 (94%)	297 (84%)	57 (16%)	2	16
2	F	354/378 (94%)	296 (84%)	58 (16%)	2	15
2	H	354/378 (94%)	297 (84%)	57 (16%)	2	16
3	I	85/85 (100%)	57 (67%)	28 (33%)	0	2
All	All	2969/3121 (95%)	2473 (83%)	496 (17%)	5	15

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	352	LYS
2	D	404	PHE
1	E	153	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 128 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	133	GLN
1	E	139	HIS
2	H	101	ASN
2	D	197	HIS
1	E	14	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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

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	-	25,30,30	3.49	8 (32%)	27,47,47	3.60	9 (33%)
5	GTP	B	500	-	27,34,34	2.38	5 (18%)	29,54,54	2.20	6 (20%)
4	GDP	C	600	-	25,30,30	3.49	8 (32%)	27,47,47	3.61	9 (33%)
5	GTP	D	500	-	27,34,34	1.40	3 (11%)	29,54,54	2.13	5 (17%)
4	GDP	E	600	-	25,30,30	3.49	8 (32%)	27,47,47	3.61	9 (33%)
5	GTP	F	500	-	27,34,34	1.42	3 (11%)	29,54,54	2.13	5 (17%)
4	GDP	G	600	-	25,30,30	3.49	8 (32%)	27,47,47	3.60	9 (33%)
5	GTP	H	500	-	27,34,34	1.41	3 (11%)	29,54,54	2.14	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 46 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	500	GTP	PG-O3B	-9.31	1.45	1.60
4	G	600	GDP	PB-O2B	-3.63	1.40	1.54
4	A	600	GDP	PB-O2B	-3.63	1.40	1.54
4	C	600	GDP	PB-O2B	-3.62	1.40	1.54
4	E	600	GDP	PB-O2B	-3.61	1.40	1.54

The worst 5 of 57 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	600	GDP	C6-C5-C4	-11.54	109.53	120.85
4	C	600	GDP	C6-C5-C4	-11.51	109.56	120.85
4	E	600	GDP	C6-C5-C4	-11.51	109.56	120.85
4	A	600	GDP	C6-C5-C4	-11.50	109.57	120.85
5	H	500	GTP	C5-C6-N1	-7.16	113.28	123.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 26 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
4	E	600	GDP	1	0
5	F	500	GTP	5	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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	H	4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Number of breaks
2	B	4
2	D	4
2	F	4

The worst 5 of 16 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	274:PRO	C	275:VAL	N	2.09
1	D	274:PRO	C	275:VAL	N	2.09
1	F	274:PRO	C	275:VAL	N	2.09
1	H	274:PRO	C	275:VAL	N	2.09
1	B	92:LEU	C	93:ILE	N	1.98