



# Full wwPDB EM Model Validation Report ⓘ

May 4, 2020 – 09:41 PM EDT

PDB ID : 6REV  
EMDB ID : EMD-4858  
Title : Cryo-EM structure of the N-terminal DC repeat (NDC) of human doublecortin (DCX) bound to 13-protofilament GDP-microtubule  
Authors : Manka, S.W.  
Deposited on : 2019-04-12  
Resolution : 3.80 Å (reported)  
Based on initial model : 1MJD

This is a Full wwPDB EM Model Validation 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.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
buster-report : 1.1.7 (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) : 2.10.1

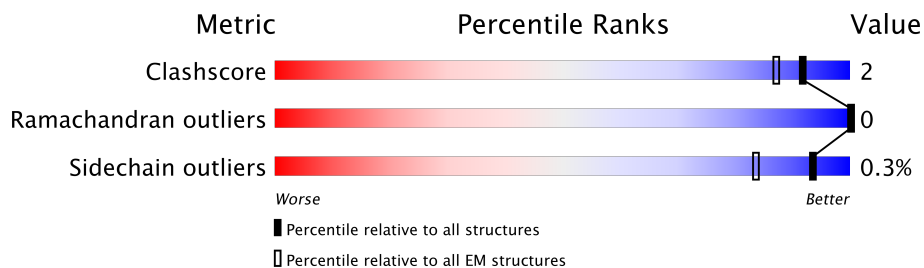
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*ELECTRON MICROSCOPY*

The reported resolution of this entry is 3.80 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	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. 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 respectively. 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	N	99	87% 12% .
2	B	429	89% 11%
2	b	429	100%
3	A	432	94% 6%
3	a	432	100%

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 14407 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 Neuronal migration protein doublecortin.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
1	N	99	785	499	140	144	2	0	0

- Molecule 2 is a protein called Tubulin beta-2B chain.

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

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
b	55	ALA	THR	conflict	UNP Q6B856
b	170	VAL	MET	conflict	UNP Q6B856
b	296	ALA	SER	conflict	UNP Q6B856
b	316	VAL	ILE	conflict	UNP Q6B856
B	55	ALA	THR	conflict	UNP Q6B856
B	170	VAL	MET	conflict	UNP Q6B856
B	296	ALA	SER	conflict	UNP Q6B856
B	316	VAL	ILE	conflict	UNP Q6B856

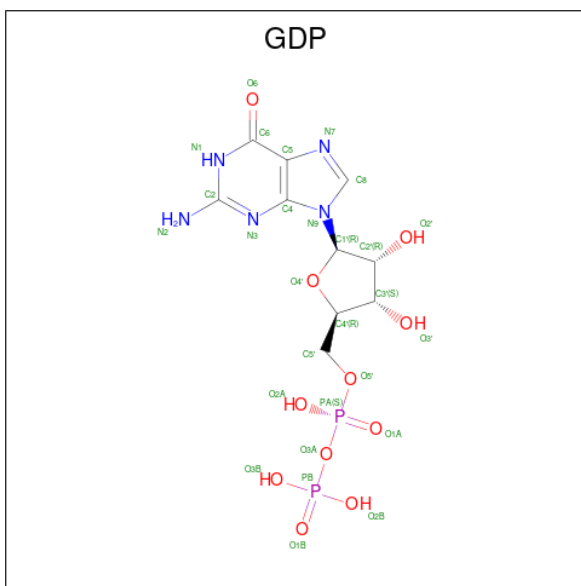
- Molecule 3 is a protein called Tubulin alpha-1B chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	A	432	3382	2144	575	642	21	0	0
3	a	432	3382	2144	575	642	21	0	0

There are 18 discrepancies between the modelled and reference sequences:

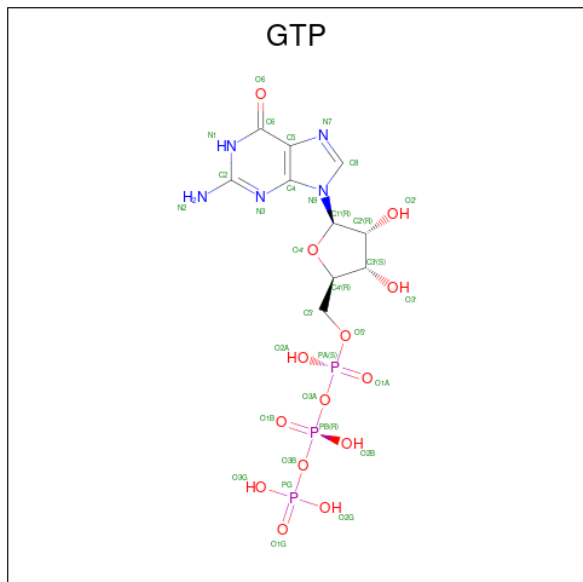
Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	SER	deletion	UNP P81947
A	?	-	ASP	deletion	UNP P81947
A	?	-	LYS	deletion	UNP P81947
A	?	-	THR	deletion	UNP P81947
A	?	-	ILE	deletion	UNP P81947
A	?	-	GLY	deletion	UNP P81947
A	?	-	GLY	deletion	UNP P81947
A	?	-	GLY	deletion	UNP P81947
A	?	-	ASP	deletion	UNP P81947
a	?	-	SER	deletion	UNP P81947
a	?	-	ASP	deletion	UNP P81947
a	?	-	LYS	deletion	UNP P81947
a	?	-	THR	deletion	UNP P81947
a	?	-	ILE	deletion	UNP P81947
a	?	-	GLY	deletion	UNP P81947
a	?	-	GLY	deletion	UNP P81947
a	?	-	GLY	deletion	UNP P81947
a	?	-	GLY	deletion	UNP P81947
a	?	-	ASP	deletion	UNP P81947

- 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	
			Total	C	N	O		P
4	b	1	28	10	5	11	2	0
4	B	1	28	10	5	11	2	0

- Molecule 5 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	
5	A	1	32	10	5	14	3	0
5	a	1	32	10	5	14	3	0


- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
6	A	1	1	1	0
6	a	1	1	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 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: Neuronal migration protein doublecortin

Chain N:  87% 12%




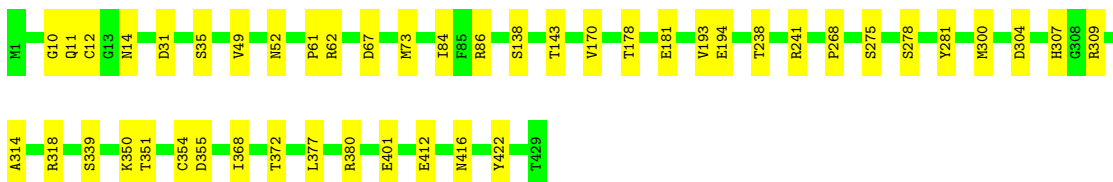
- Molecule 2: Tubulin beta-2B chain

Chain b:  100%

There are no outlier residues recorded for this chain.

- Molecule 2: Tubulin beta-2B chain

Chain B:  89% 11%



- Molecule 3: Tubulin alpha-1B chain

Chain A:  94% 6%



- Molecule 3: Tubulin alpha-1B chain

Chain a:  100%

There are no outlier residues recorded for this chain.

## 4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	30434	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI POLARA 300	Depositor
Voltage (kV)	300	Depositor
Electron dose ( $e^-/\text{\AA}^2$ )	25	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (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: GDP, 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  >5	RMSZ	# Z  >5
1	N	0.31	0/798	0.55	1/1069 (0.1%)
2	B	0.30	0/3443	0.49	0/4666
2	b	0.30	0/3443	0.49	0/4666
3	A	0.31	0/3459	0.49	0/4697
3	a	0.31	0/3459	0.49	0/4697
All	All	0.30	0/14602	0.50	1/19795 (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	N	108	ASP	C-N-CA	5.01	132.82	122.30

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	N	785	0	776	6	0
2	B	3368	0	3246	25	0
2	b	3368	0	3246	0	0
3	A	3382	0	3292	16	0
3	a	3382	0	3292	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	28	0	12	1	0
4	b	28	0	12	0	0
5	A	32	0	12	2	0
5	a	32	0	12	0	0
6	A	1	0	0	0	0
6	a	1	0	0	0	0
All	All	14407	0	13900	44	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 2.

All (44) 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)
3:A:249:ASN:H	2:B:11:GLN:HE22	1.46	0.63
1:N:76:ARG:NH2	2:B:401:GLU:OE2	2.31	0.63
3:A:71:GLU:HB3	3:A:98:ASP:HB3	1.82	0.62
3:A:5:ILE:HG22	3:A:64:ARG:HB3	1.85	0.59
2:B:354:CYS:SG	2:B:355:ASP:N	2.78	0.56
1:N:53:LYS:HG2	1:N:121:GLU:HG3	1.88	0.56
2:B:178:THR:HB	2:B:181:GLU:HB3	1.88	0.55
3:A:259:LEU:HD11	3:A:316:CYS:HB2	1.87	0.55
2:B:309:ARG:NH1	2:B:339:SER:O	2.40	0.55
2:B:412:GLU:OE2	2:B:416:ASN:ND2	2.41	0.54
2:B:372:THR:HG22	2:B:422:TYR:HD1	1.73	0.53
3:A:215:ARG:NH2	3:A:299:ALA:O	2.43	0.51
3:A:144:GLY:N	5:A:501:GTP:O2B	2.44	0.50
3:A:195:LEU:HD12	3:A:266:HIS:HE1	1.77	0.50
2:B:143:THR:N	4:B:501:GDP:O1B	2.44	0.50
2:B:304:ASP:HB3	2:B:307:HIS:CD2	2.47	0.49
1:N:95:ILE:HG23	2:B:380:ARG:HH12	57.51	0.48
2:B:49:VAL:HG11	2:B:241:ARG:HG2	1.96	0.48
3:A:103:TYR:HB2	3:A:186:ASN:HD22	1.77	0.48
3:A:71:GLU:OE2	5:A:501:GTP:O2G	2.32	0.48
2:B:314:ALA:HB3	2:B:368:ILE:HB	1.95	0.47
2:B:31:ASP:OD1	2:B:35:SER:N	2.47	0.47
3:A:64:ARG:NH2	3:A:129:CYS:SG	2.88	0.47
2:B:170:VAL:HG11	2:B:377:LEU:HD21	1.97	0.47
2:B:86:ARG:HG2	2:B:281:TYR:HB3	46.57	0.47
2:B:52:ASN:OD1	2:B:62:ARG:NH2	2.44	0.47
1:N:59:ARG:HD2	1:N:96:ASN:HD22	1.79	0.47

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:12:CYS:HG	2:B:138:SER:HG	1.62	0.46
3:A:238:ILE:HG12	3:A:378:LEU:HD11	1.98	0.46
1:N:84:LEU:HD13	1:N:87:LEU:HD12	1.97	0.46
2:B:67:ASP:HB2	2:B:73:MET:HE2	1.98	0.46
2:B:61:PRO:HD3	2:B:84:ILE:HG12	1.99	0.45
1:N:51:LYS:NZ	1:N:121:GLU:OE2	2.47	0.45
3:A:151:SER:HG	3:A:190:THR:HG1	1.64	0.45
3:A:62:VAL:HG11	3:A:283:HIS:HA	45.25	0.44
2:B:10:GLY:O	2:B:14:ASN:ND2	2.51	0.44
2:B:268:PRO:HG2	2:B:300:MET:HB2	2.00	0.43
3:A:180:ALA:HB3	3:A:183:GLU:HG3	2.01	0.42
3:A:27:GLU:OE1	3:A:243:ARG:NH1	2.40	0.42
2:B:350:LYS:NZ	2:B:351:THR:O	2.52	0.42
2:B:238:THR:HG21	2:B:318:ARG:HD2	2.02	0.41
2:B:275:SER:HB3	2:B:278:SER:HB3	2.03	0.40
2:B:193:VAL:HG23	2:B:194:GLU:HG2	2.03	0.40
3:A:136:LEU:HD23	3:A:167:LEU:HB2	2.04	0.40

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	N	97/99 (98%)	83 (86%)	14 (14%)	0	100	100
2	B	427/429 (100%)	405 (95%)	22 (5%)	0	100	100
2	b	427/429 (100%)	405 (95%)	22 (5%)	0	100	100
3	A	428/432 (99%)	406 (95%)	22 (5%)	0	100	100
3	a	428/432 (99%)	406 (95%)	22 (5%)	0	100	100
All	All	1807/1821 (99%)	1705 (94%)	102 (6%)	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	N	82/88 (93%)	78 (95%)	4 (5%)	27	62
2	B	368/369 (100%)	368 (100%)	0	100	100
2	b	368/369 (100%)	368 (100%)	0	100	100
3	A	364/366 (100%)	364 (100%)	0	100	100
3	a	364/366 (100%)	364 (100%)	0	100	100
All	All	1546/1558 (99%)	1542 (100%)	4 (0%)	93	97

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	N	51	LYS
1	N	66	LYS
1	N	89	ARG
1	N	131	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (29) such sidechains are listed below:

Mol	Chain	Res	Type
1	N	131	ASN
2	b	8	GLN
2	b	11	GLN
2	b	137	HIS
2	b	190	HIS
2	b	226	ASN
2	b	292	GLN
2	b	396	HIS
3	A	28	HIS
3	A	88	HIS
3	A	107	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	A	186	ASN
3	A	192	HIS
3	A	206	ASN
3	A	266	HIS
2	B	8	GLN
2	B	11	GLN
2	B	137	HIS
2	B	190	HIS
2	B	226	ASN
2	B	307	HIS
3	a	28	HIS
3	a	88	HIS
3	a	107	HIS
3	a	186	ASN
3	a	192	HIS
3	a	206	ASN
3	a	266	HIS
3	a	283	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 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$
4	GDP	B	501	-	24,30,30	0.92	1 (4%)	31,47,47	2.03	5 (16%)
5	GTP	A	501	6	26,34,34	0.89	1 (3%)	33,54,54	1.98	5 (15%)
5	GTP	a	501	6	26,34,34	0.89	1 (3%)	33,54,54	1.97	5 (15%)
4	GDP	b	501	-	24,30,30	0.93	1 (4%)	31,47,47	2.02	5 (16%)

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	B	501	-	-	2/12/32/32	0/3/3/3
5	GTP	A	501	6	-	3/18/38/38	0/3/3/3
5	GTP	a	501	6	-	3/18/38/38	0/3/3/3
4	GDP	b	501	-	-	2/12/32/32	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	b	501	GDP	C6-N1	3.01	1.38	1.33
4	B	501	GDP	C6-N1	3.01	1.38	1.33
5	A	501	GTP	C6-N1	2.97	1.38	1.33
5	a	501	GTP	C6-N1	2.97	1.38	1.33

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	a	501	GTP	C5-C6-N1	-8.16	112.27	123.43
5	A	501	GTP	C5-C6-N1	-8.14	112.29	123.43
4	B	501	GDP	C5-C6-N1	-8.12	112.32	123.43
4	b	501	GDP	C5-C6-N1	-8.08	112.38	123.43
4	B	501	GDP	C6-N1-C2	5.80	125.14	115.93
4	b	501	GDP	C6-N1-C2	5.73	125.04	115.93
5	A	501	GTP	C6-N1-C2	5.73	125.04	115.93
5	a	501	GTP	C6-N1-C2	5.73	125.04	115.93
4	B	501	GDP	N3-C2-N1	-3.04	123.16	127.22
4	b	501	GDP	N3-C2-N1	-2.97	123.26	127.22

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	501	GTP	N3-C2-N1	-2.97	123.26	127.22
5	a	501	GTP	N3-C2-N1	-2.95	123.29	127.22
5	A	501	GTP	C6-C5-C4	-2.47	118.44	120.80
4	b	501	GDP	C6-C5-C4	-2.46	118.45	120.80
4	B	501	GDP	C6-C5-C4	-2.45	118.46	120.80
5	a	501	GTP	C6-C5-C4	-2.43	118.48	120.80
4	b	501	GDP	C2-N3-C4	-2.31	112.72	115.36
4	B	501	GDP	C2-N3-C4	-2.27	112.76	115.36
5	A	501	GTP	C2-N3-C4	-2.21	112.83	115.36
5	a	501	GTP	C2-N3-C4	-2.20	112.84	115.36

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	501	GTP	C5'-O5'-PA-O1A
5	a	501	GTP	C5'-O5'-PA-O1A
4	b	501	GDP	PB-O3A-PA-O1A
4	B	501	GDP	PB-O3A-PA-O1A
4	b	501	GDP	C4'-C5'-O5'-PA
4	B	501	GDP	C4'-C5'-O5'-PA
5	A	501	GTP	C5'-O5'-PA-O3A
5	a	501	GTP	C5'-O5'-PA-O3A
5	A	501	GTP	C5'-O5'-PA-O2A
5	a	501	GTP	C5'-O5'-PA-O2A

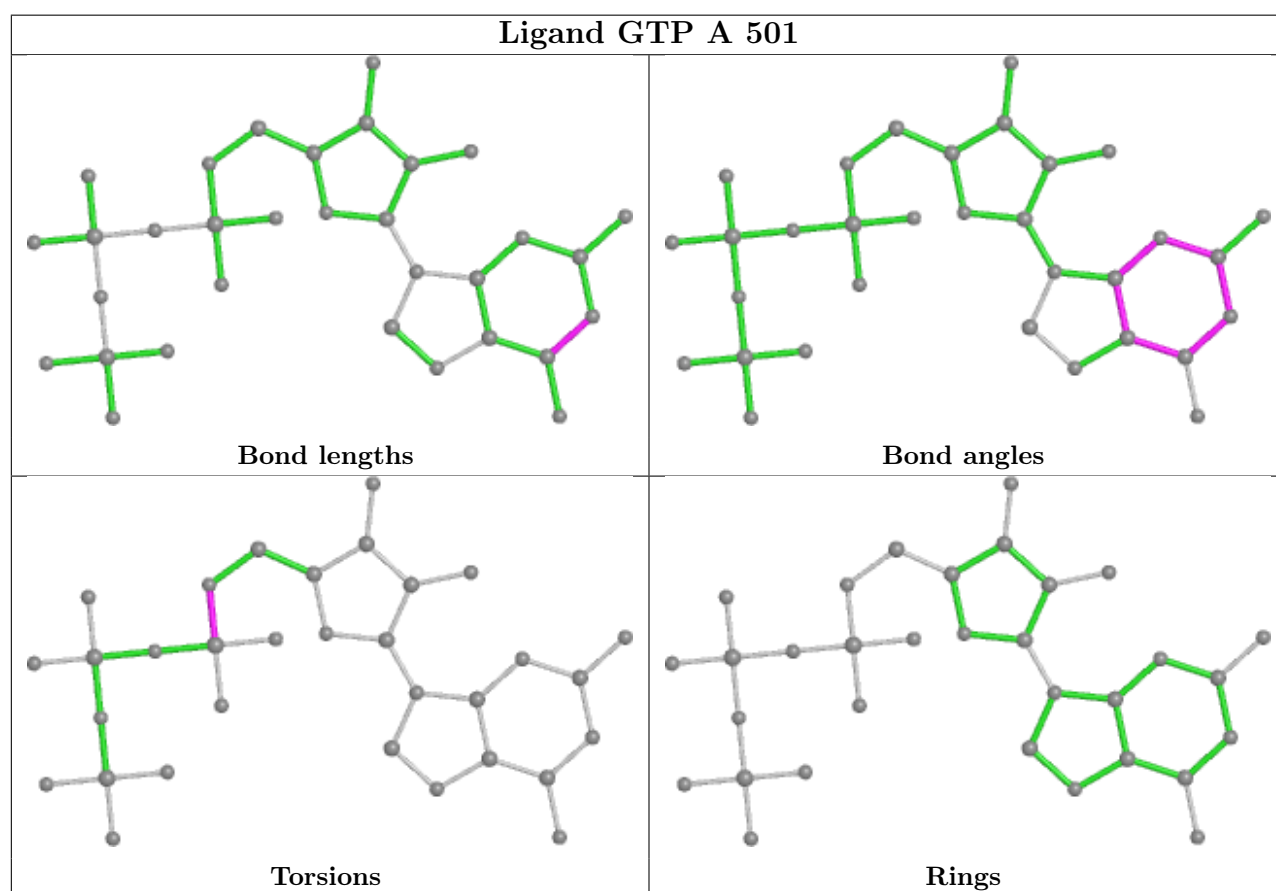
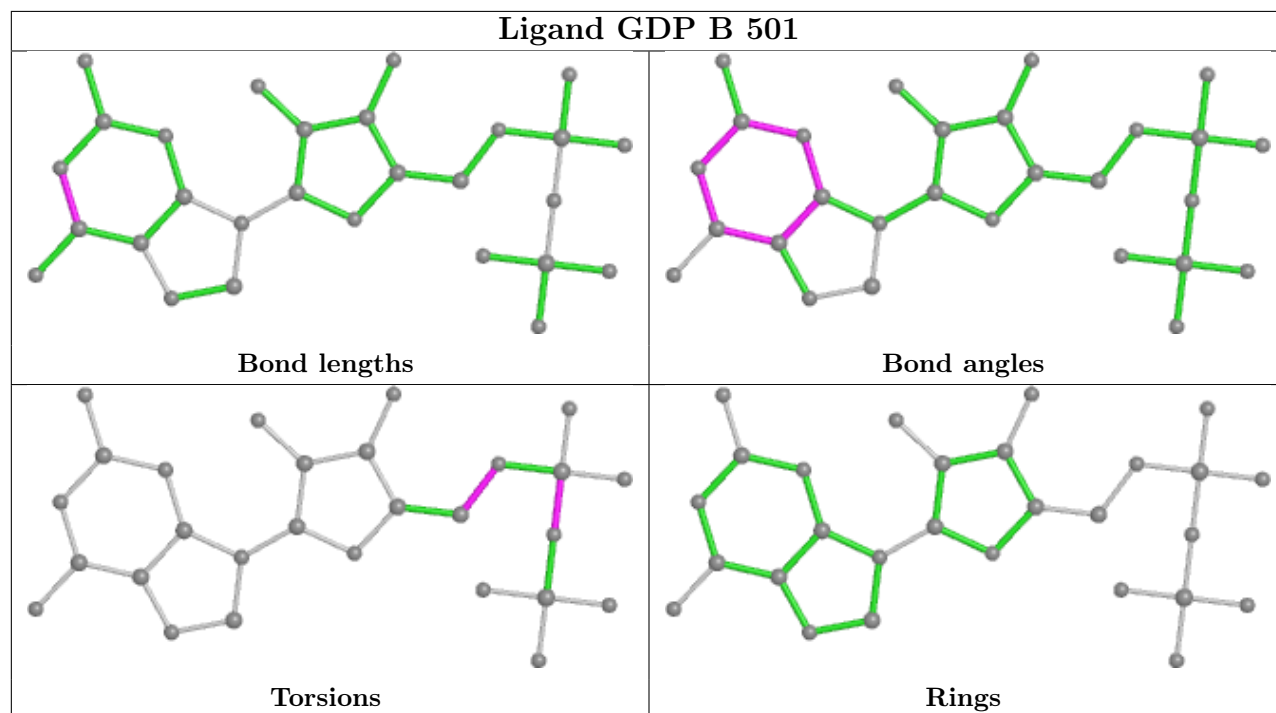
There are no ring outliers.

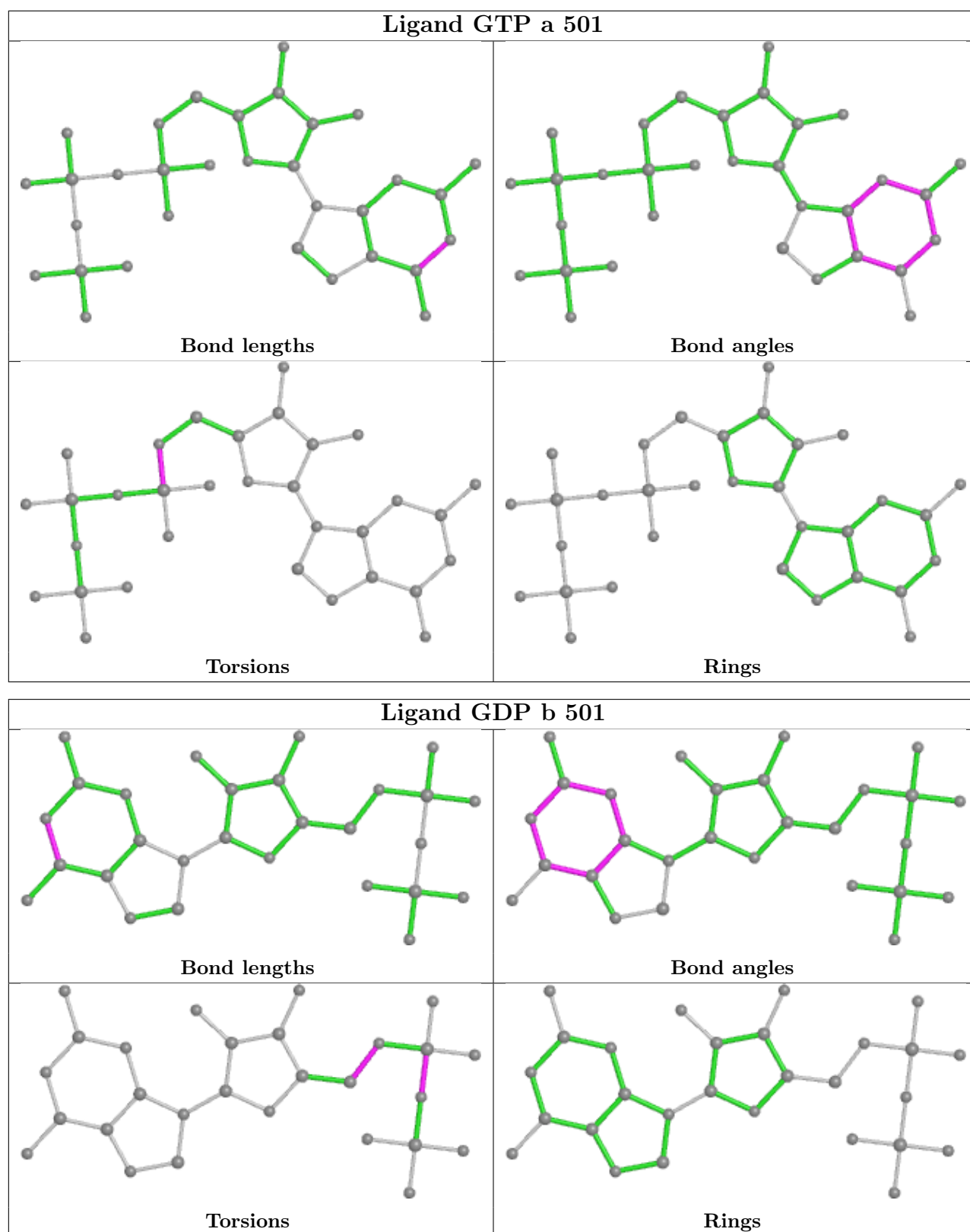
2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	501	GDP	1	0
5	A	501	GTP	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and

any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





## 5.7 Other polymers [i](#)

There are no such residues in this entry.



## 5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
3	A	1
3	a	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	37:PRO	C	47:ASP	N	11.53
1	a	37:PRO	C	47:ASP	N	11.53