wwPDB EM Validation Summary Report

May 27, 2024 – 02:56 PM JST

PDB ID	:	2HXF
Title	:	KIF1A head-microtubule complex structure in amppnp-form
Authors	:	Kikkawa, M.; Hirokawa, N.
Deposited on	:	2006-08-03
Resolution	:	10.00 Å(reported)
Based on initial models	:	1VFV, 1JFF

This is a wwPDB EM 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
https://www.wwpdb.org/validation/2017/EMValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

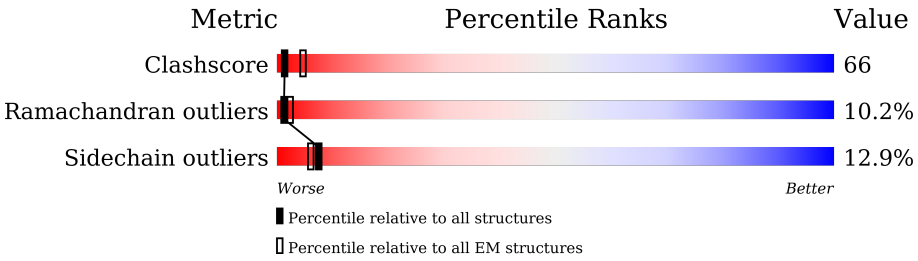
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

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 10.00 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	158937	4297
Ramachandran outliers	154571	4023
Sidechain outliers	154315	3826

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 of the bar indicate the fraction of residues that contain outliers for >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	451	16% 56% 18% • 9%
2	B	445	20% 56% 18% • •
3	C	394	57% 21% • • 18%

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 9279 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
1	A	412	Total	C	N	O	S	0	0
			3227	2043	551	613	20		

- Molecule 2 is a protein called Tubulin beta chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	426	Total	C	N	O	S	0	0
			3351	2105	575	646	25		

- Molecule 3 is a protein called Kinesin-like protein KIF1A.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	324	Total	C	N	O	S	0	0
			2546	1581	446	505	14		

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-15	MET	-	cloning artifact	UNP P33173
C	-14	ALA	-	cloning artifact	UNP P33173
C	-13	SER	-	cloning artifact	UNP P33173
C	-12	MET	-	cloning artifact	UNP P33173
C	-11	THR	-	cloning artifact	UNP P33173
C	-10	GLY	-	cloning artifact	UNP P33173
C	-9	GLY	-	cloning artifact	UNP P33173
C	-8	GLN	-	cloning artifact	UNP P33173
C	-7	GLN	-	cloning artifact	UNP P33173
C	-6	MET	-	cloning artifact	UNP P33173
C	-5	GLY	-	cloning artifact	UNP P33173
C	-4	ARG	-	cloning artifact	UNP P33173
C	-3	ASP	-	cloning artifact	UNP P33173
C	-2	PRO	-	cloning artifact	UNP P33173
C	-1	ILE	-	cloning artifact	UNP P33173

Continued on next page...

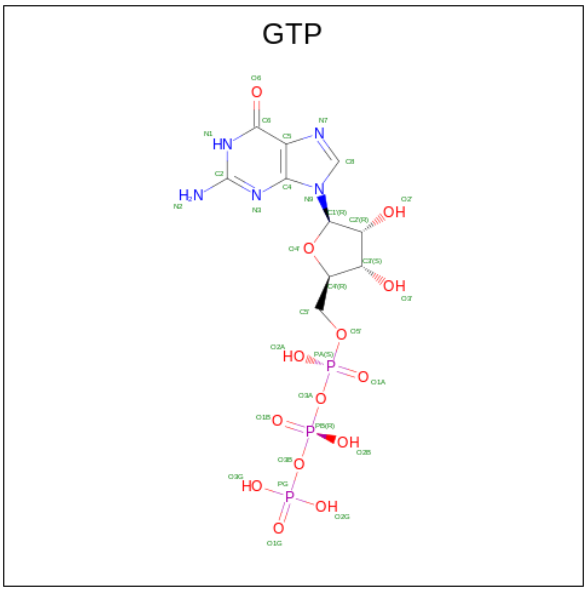
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	ASN	-	cloning artifact	UNP P33173
C	1	MET	-	cloning artifact	UNP P33173
C	2	PRO	-	cloning artifact	UNP P33173
C	202	ALA	PRO	engineered mutation	UNP P33173
C	356	ASN	-	linker	UNP P33173
C	357	THR	-	linker	UNP P33173
C	358	VAL	-	linker	UNP P33173
C	359	SER	-	linker	UNP P33173
C	360	VAL	-	linker	UNP P33173
C	361	ASN	-	linker	UNP P33173
C	362	LEU	-	linker	UNP P33173
C	363	GLU	-	linker	UNP P33173
C	364	LEU	-	linker	UNP P33173
C	365	THR	-	linker	UNP P33173
C	366	ALA	-	linker	UNP P33173
C	367	GLU	-	linker	UNP P33173
C	368	GLU	-	linker	UNP P33173
C	369	TRP	-	linker	UNP P33173
C	370	LYS	-	linker	UNP P33173
C	371	LYS	-	linker	UNP P33173
C	372	LYS	-	linker	UNP P33173
C	373	HIS	-	expression tag	UNP P33173
C	374	HIS	-	expression tag	UNP P33173
C	375	HIS	-	expression tag	UNP P33173
C	376	HIS	-	expression tag	UNP P33173
C	377	HIS	-	expression tag	UNP P33173
C	378	HIS	-	expression tag	UNP P33173

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

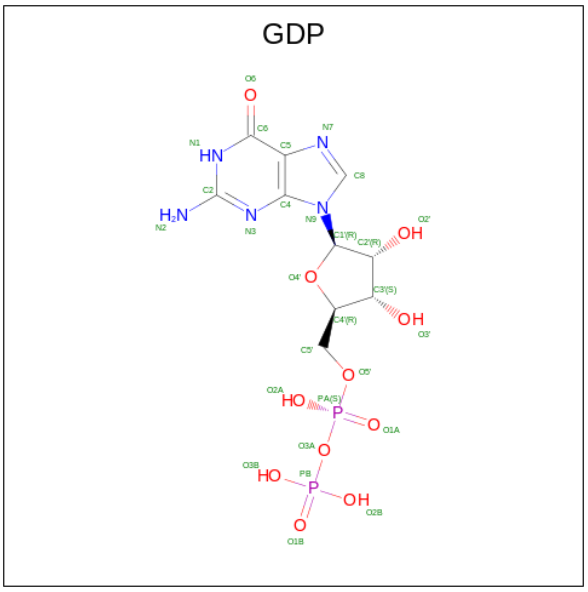
Mol	Chain	Residues	Atoms	AltConf
4	A	1	Total Mg 1 1	0
4	C	1	Total Mg 1 1	0

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



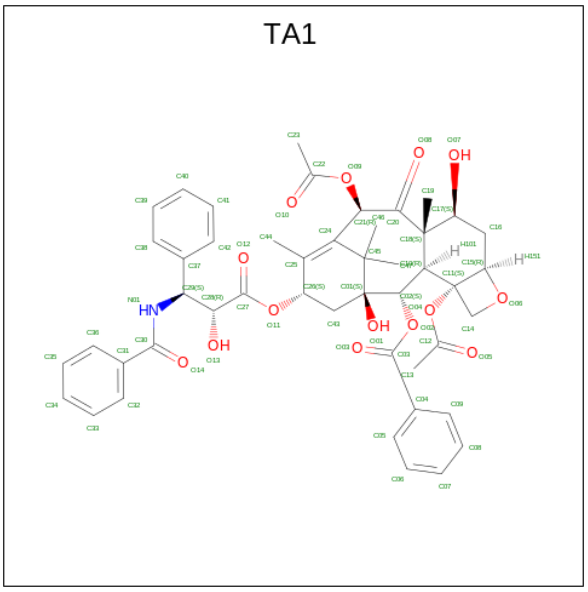
Mol	Chain	Residues	Atoms					AltConf
5	A	1	Total	C	N	O	P	0
			32	10	5	14	3	

- Molecule 6 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).



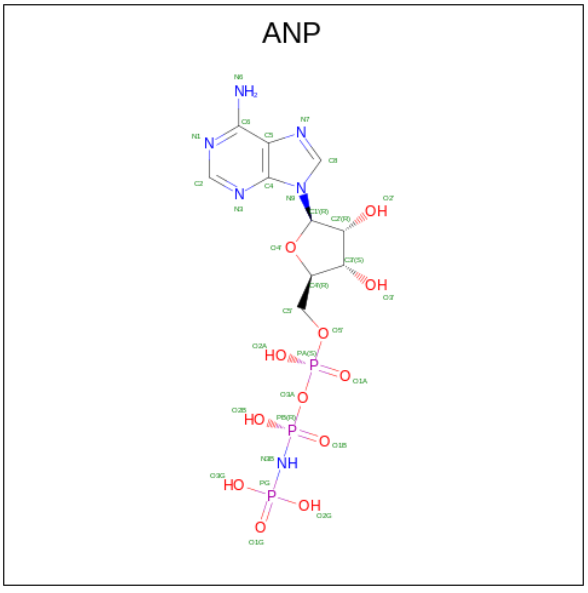
Mol	Chain	Residues	Atoms					AltConf
6	B	1	Total	C	N	O	P	0
			28	10	5	11	2	

- Molecule 7 is TAXOL (three-letter code: TA1) (formula: $C_{47}H_{51}NO_{14}$).



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

- Molecule 8 is PHOSPHOAMINOPHOSPHONIC ACID-ADENYLATE ESTER (three-letter code: ANP) (formula: $C_{10}H_{17}N_6O_{12}P_3$).



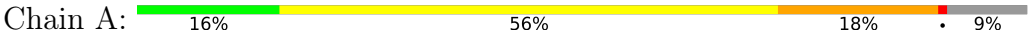
Mol	Chain	Residues	Atoms					AltConf
8	C	1	Total	C	N	O	P	0
			31	10	6	12	3	

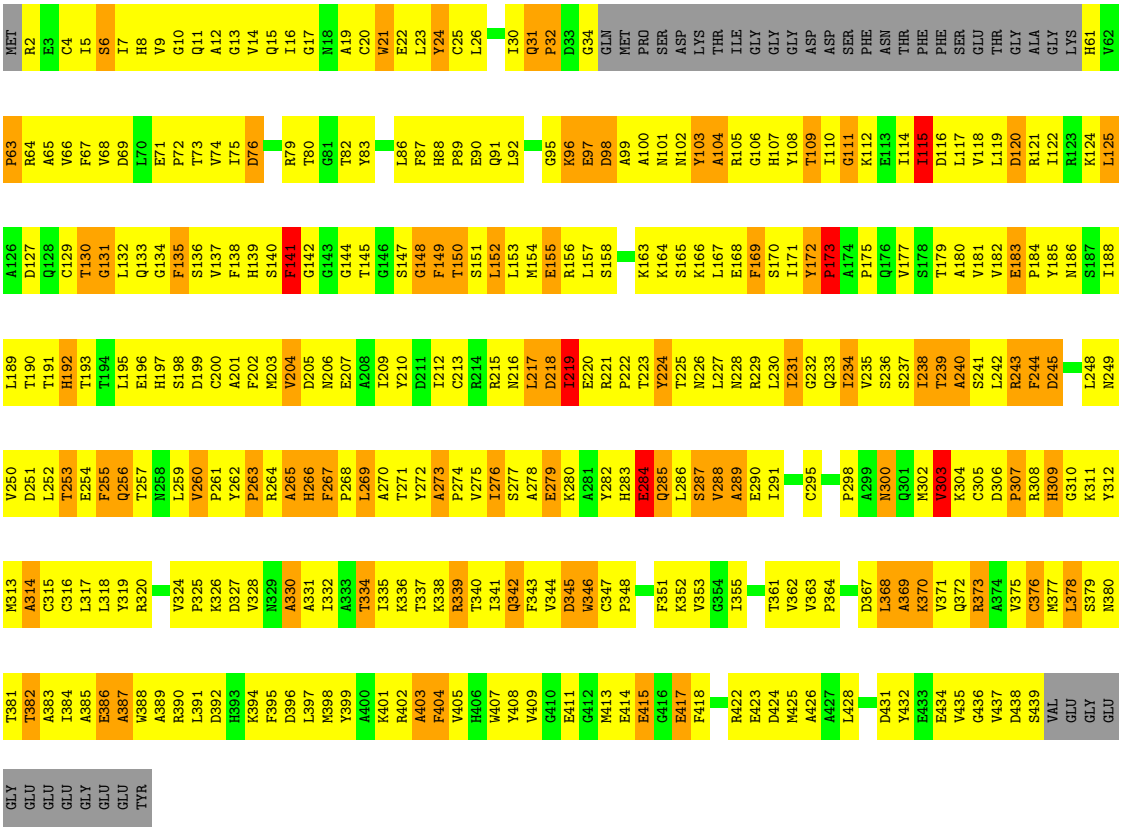
3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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 alpha chain

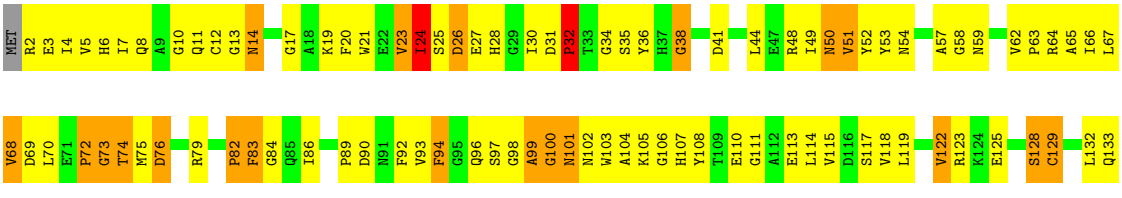
Chain A:

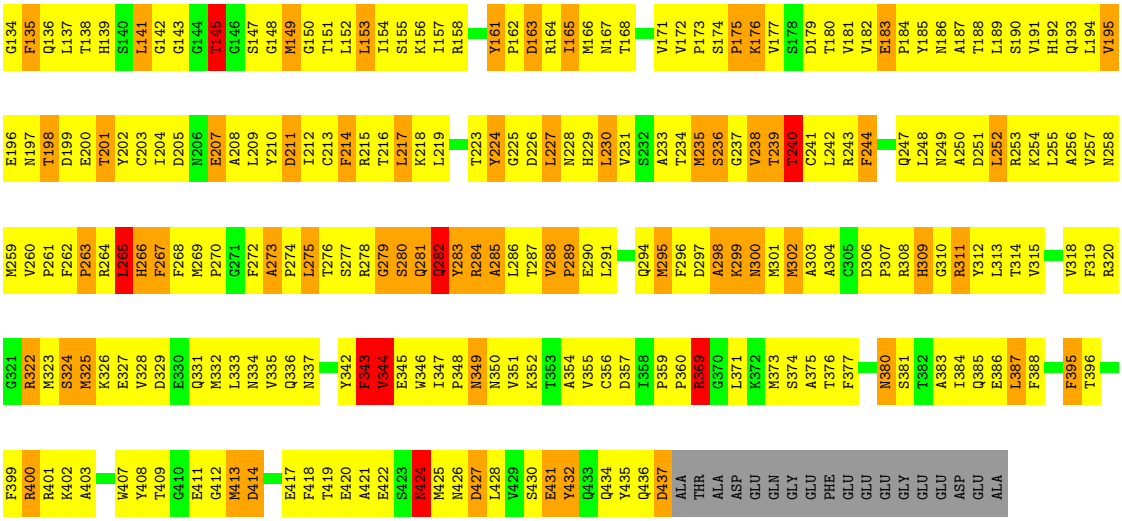




- Molecule 2: Tubulin beta chain

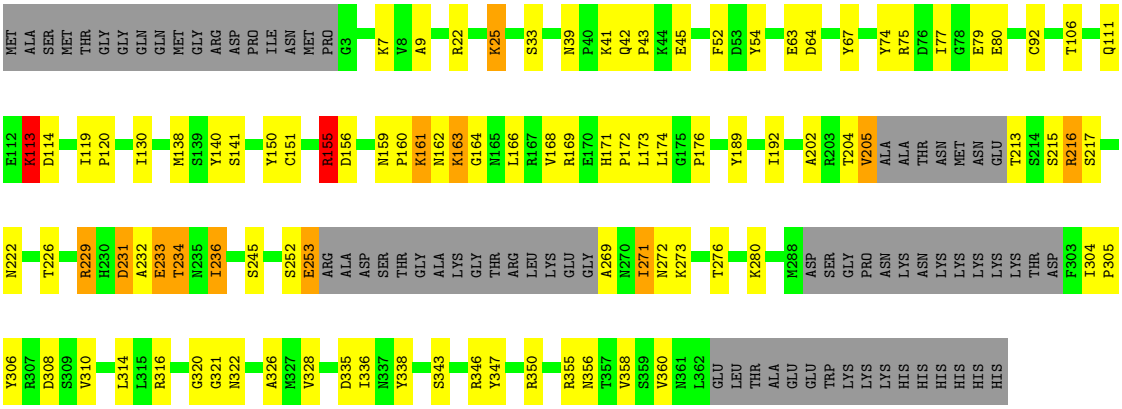
Chain B:





• Molecule 3: Kinesin-like protein KIF1A

Chain C:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS failed to run properly - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	1.00Å 1.00Å 1.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	(Not available) – 10.00	Depositor
% Data completeness (in resolution range)	(Not available) ((Not available)-10.00)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	unknown	Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	9279	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

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: MG, GTP, ANP, TA1, GDP

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	A	0.50	0/3300	0.73	0/4482
2	B	0.51	0/3426	0.76	2/4642 (0.0%)
3	C	0.46	0/2585	0.80	3/3491 (0.1%)
All	All	0.49	0/9311	0.76	5/12615 (0.0%)

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	C	0	2

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	213	THR	CA-CB-CG2	-13.59	93.38	112.40
3	C	155	ARG	NE-CZ-NH2	-8.47	116.07	120.30
2	B	235	MET	CG-SD-CE	6.09	109.95	100.20
3	C	155	ARG	CD-NE-CZ	-6.03	115.16	123.60
2	B	217	LEU	N-CA-C	-5.36	96.53	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	C	155	ARG	Sidechain
3	C	74	TYR	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	3227	0	3141	566	0
2	B	3351	0	3227	569	0
3	C	2546	0	2507	132	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
5	A	32	0	12	5	0
6	B	28	0	12	1	0
7	B	62	0	51	5	0
8	C	31	0	13	0	0
All	All	9279	0	8963	1206	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 66.

The worst 5 of 1206 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:C:204:THR:CA	3:C:205:VAL:HB	1.63	1.27
1:A:414:GLU:HB2	3:C:253:GLU:O	1.32	1.25
1:A:409:VAL:HG12	3:C:272:ASN:CB	1.69	1.23
2:B:234:THR:HG21	2:B:270:PRO:HB2	1.23	1.17
1:A:243:ARG:NH2	1:A:252:LEU:H	1.45	1.14

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	408/451 (90%)	266 (65%)	83 (20%)	59 (14%)	0 4
2	B	424/445 (95%)	273 (64%)	95 (22%)	56 (13%)	0 5
3	C	316/394 (80%)	309 (98%)	5 (2%)	2 (1%)	25 66
All	All	1148/1290 (89%)	848 (74%)	183 (16%)	117 (10%)	1 9

5 of 117 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	96	LYS
1	A	97	GLU
1	A	108	TYR
1	A	109	THR
1	A	141	PHE

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	347/377 (92%)	298 (86%)	49 (14%)	3 16
2	B	367/381 (96%)	307 (84%)	60 (16%)	2 13
3	C	286/345 (83%)	266 (93%)	20 (7%)	15 40
All	All	1000/1103 (91%)	871 (87%)	129 (13%)	7 18

5 of 129 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	161	LYS
3	C	229	ARG
1	A	432	TYR
1	A	431	ASP
3	C	234	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	101	ASN
3	C	42	GLN
2	B	136	GLN
3	C	171	HIS
2	B	380	ASN

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 monosaccharides 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	$\#\|Z\| > 2$
6	GDP	B	600	-	24,30,30	2.60	9 (37%)	30,47,47	2.93	8 (26%)
5	GTP	A	500	4	26,34,34	1.29	4 (15%)	32,54,54	1.10	3 (9%)
7	TA1	B	601	-	68,68,68	2.00	19 (27%)	105,105,105	1.39	11 (10%)
8	ANP	C	1500	4	29,33,33	1.93	9 (31%)	31,52,52	3.51	14 (45%)

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
6	GDP	B	600	-	-	4/12/32/32	0/3/3/3
5	GTP	A	500	4	-	3/18/38/38	0/3/3/3
7	TA1	B	601	-	-	9/41/127/127	0/7/7/7
8	ANP	C	1500	4	-	9/14/38/38	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)
6	B	600	GDP	O4'-C1'	6.27	1.49	1.41
6	B	600	GDP	O6-C6	5.72	1.34	1.23
7	B	601	TA1	C06-C05	5.29	1.50	1.38
7	B	601	TA1	C18-C10	5.10	1.69	1.57
8	C	1500	ANP	PB-O3A	4.93	1.65	1.59

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	B	600	GDP	C8-N7-C5	9.29	120.68	102.99
8	C	1500	ANP	O2B-PB-O1B	9.14	129.08	109.92
8	C	1500	ANP	O5'-PA-O1A	-8.10	77.41	109.07
8	C	1500	ANP	O2B-PB-O3A	-7.60	79.30	104.64
6	B	600	GDP	N2-C2-N3	6.28	131.97	119.74

There are no chirality outliers.

5 of 25 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	B	600	GDP	PA-O3A-PB-O2B
6	B	600	GDP	C5'-O5'-PA-O3A
6	B	600	GDP	C5'-O5'-PA-O1A
8	C	1500	ANP	PG-N3B-PB-O1B
8	C	1500	ANP	PA-O3A-PB-O1B

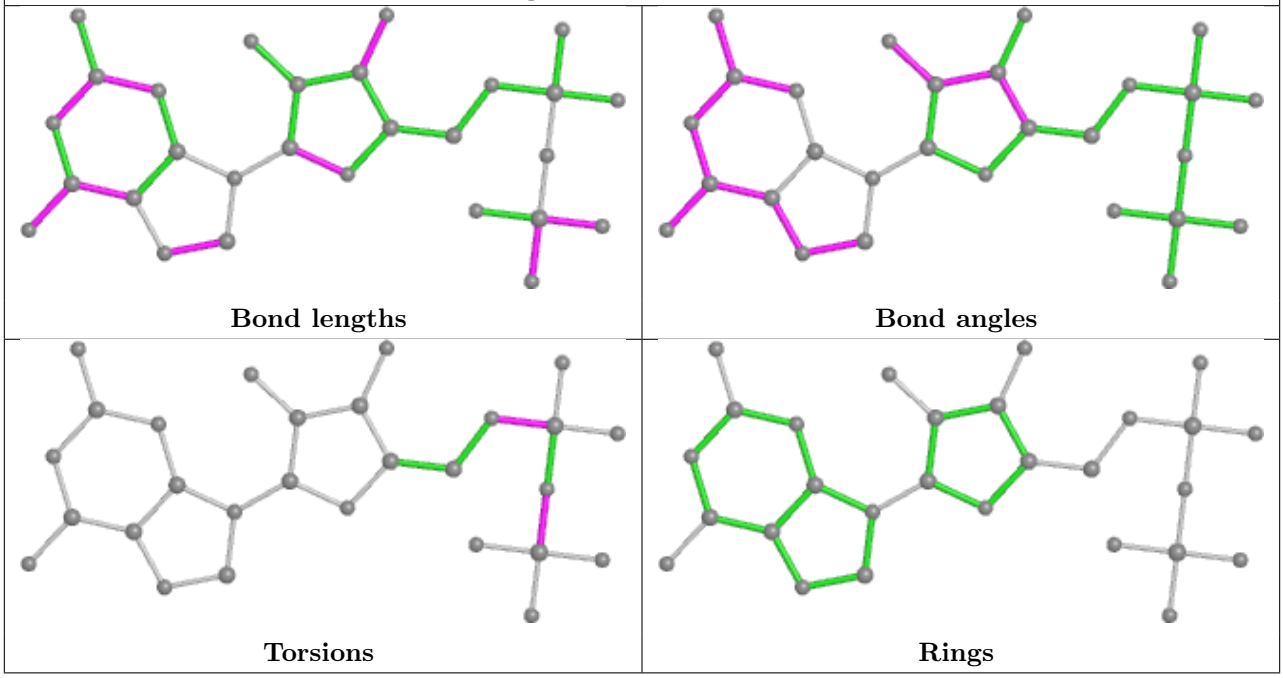
There are no ring outliers.

3 monomers are involved in 11 short contacts:

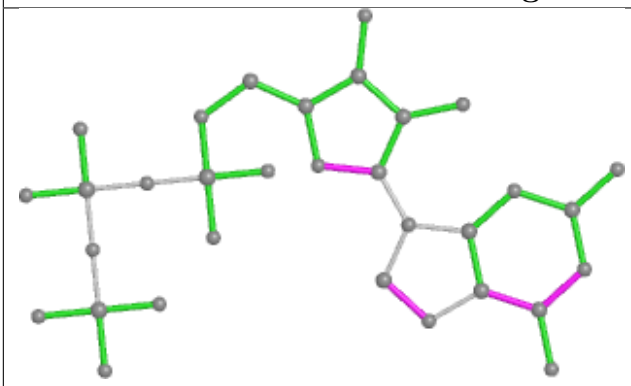
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	600	GDP	1	0
5	A	500	GTP	5	0
7	B	601	TA1	5	0

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 then 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.

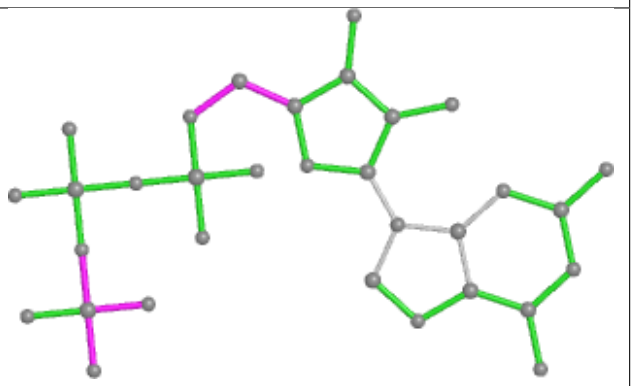
Ligand GDP B 600



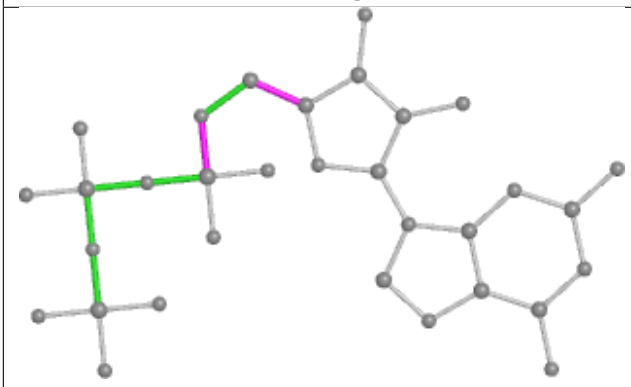
Ligand GTP A 500



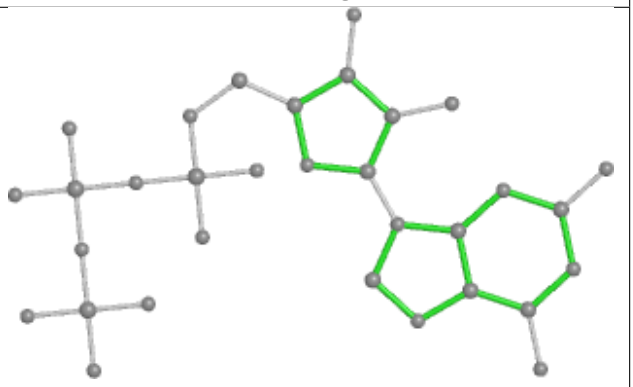
Bond lengths



Bond angles

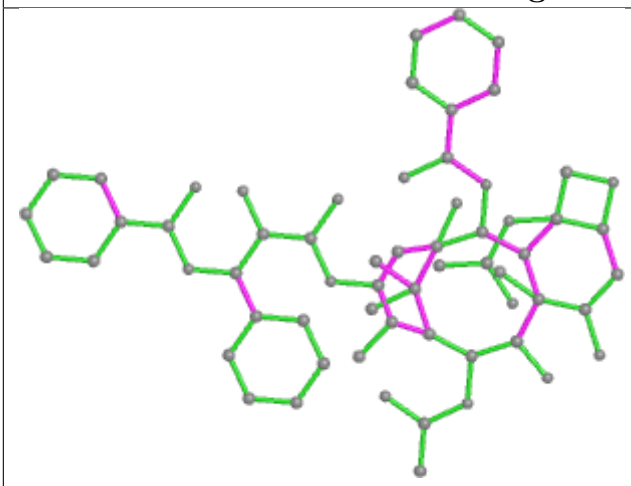


Torsions

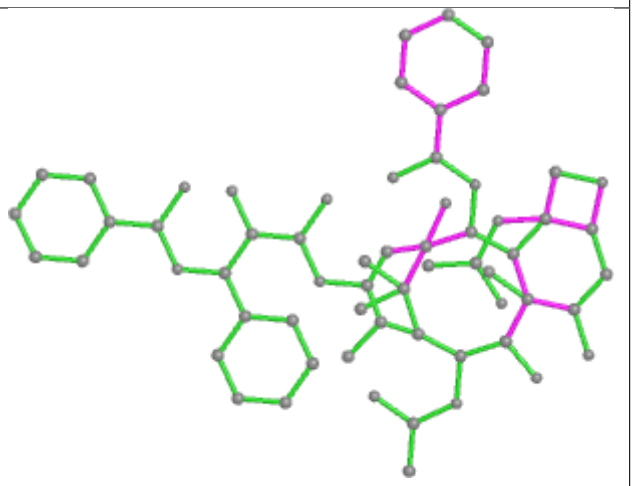


Rings

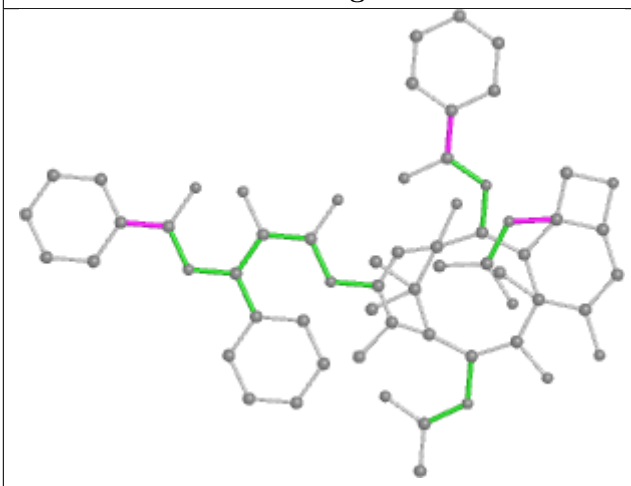
Ligand TA1 B 601



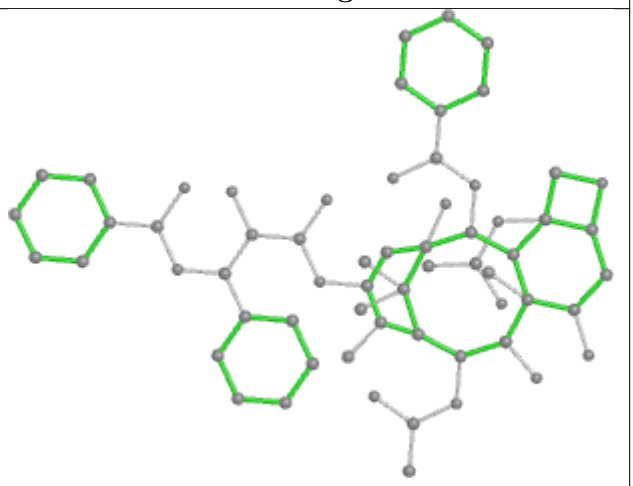
Bond lengths



Bond angles

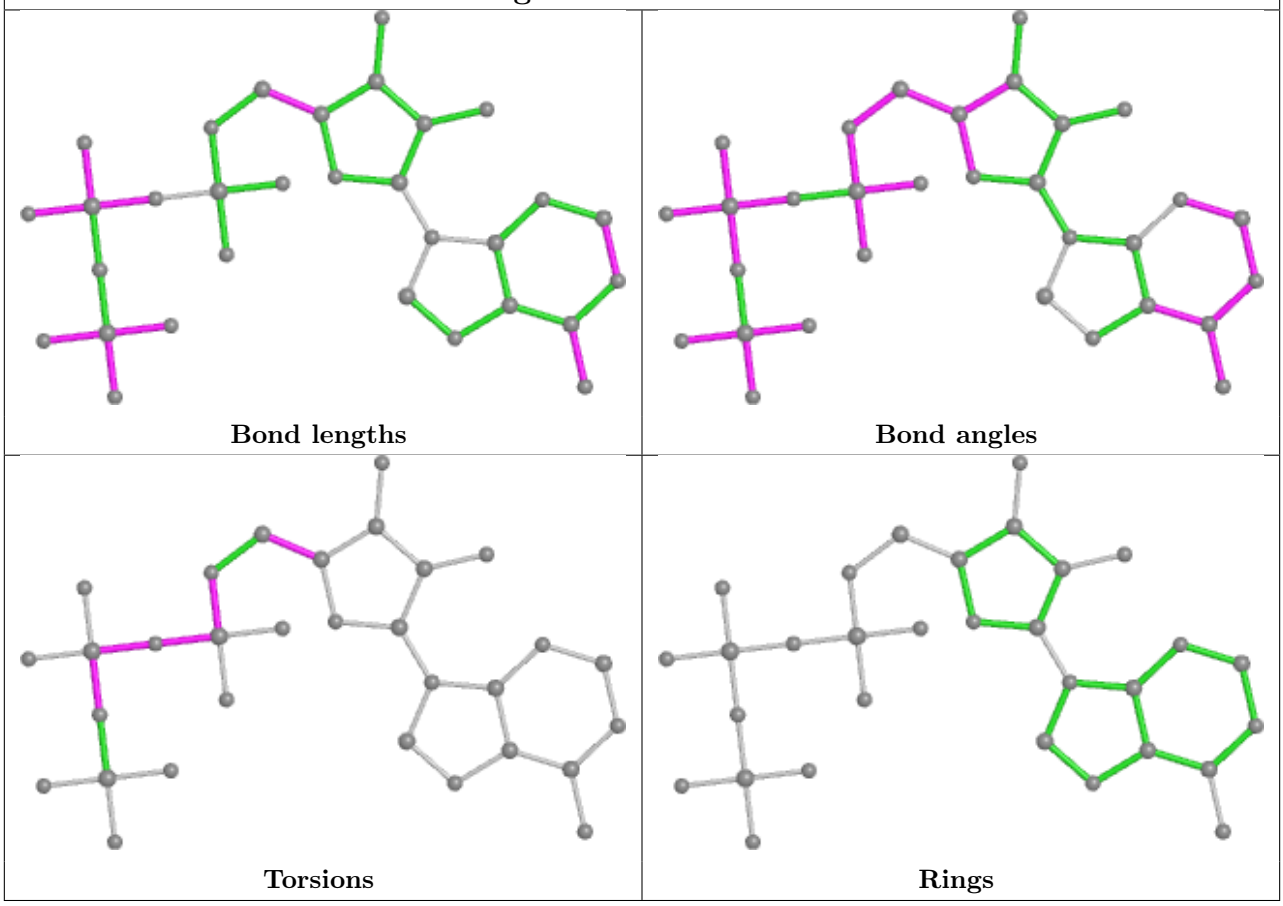


Torsions



Rings

Ligand ANP C 1500



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.