wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:36 PM BST

PDB ID	:	1IA0
Title	:	KIF1A HEAD-MICROTUBULE COMPLEX STRUCTURE IN ATP-FORM
Authors	:	Kikkawa, M.; Sablin, E.P.; Okada, Y.; Yajima, H.; Fletterick, R.J.; Hirokawa, N.
Deposited on	:	2001-03-22
Resolution	:	15.00 Å(reported)

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

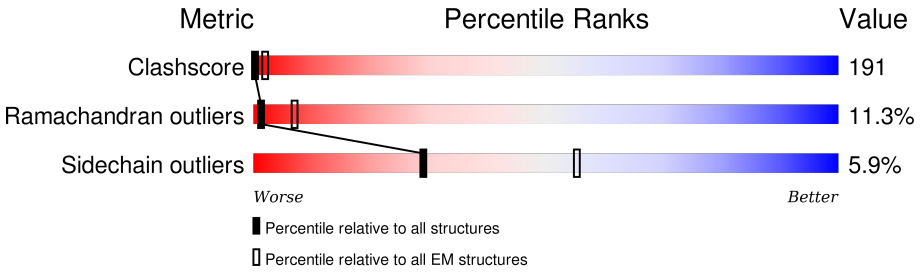
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 15.00 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	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	7% 51% 26% 14% •
2	B	445	5% 47% 28% 16% •
3	K	394	64% 16% • 17%

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
5	GTP	A	500	-	-	X	-
6	GDP	B	501	-	-	X	-
7	TXL	B	502	-	-	X	-

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 9606 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	440	Total	C	N	O	S	0	0
			3430	2168	583	657	22		

- Molecule 2 is a protein called TUBULIN BETA CHAIN.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	427	Total	C	N	O	S	0	0
			3359	2110	576	647	26		

- Molecule 3 is a protein called KINESIN-LIKE PROTEIN KIF1A.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	K	328	Total	C	N	O	S	10	0
			2667	1651	469	531	16		

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	-15	MET	-	see remark 999	UNP P33173
K	-14	ALA	-	see remark 999	UNP P33173
K	-13	SER	-	see remark 999	UNP P33173
K	-12	MET	-	see remark 999	UNP P33173
K	-11	THR	-	see remark 999	UNP P33173
K	-10	GLY	-	see remark 999	UNP P33173
K	-9	GLY	-	see remark 999	UNP P33173
K	-8	GLN	-	see remark 999	UNP P33173
K	-7	GLN	-	see remark 999	UNP P33173
K	-6	MET	-	see remark 999	UNP P33173
K	-5	GLY	-	see remark 999	UNP P33173
K	-4	ARG	-	see remark 999	UNP P33173
K	-3	ASP	-	see remark 999	UNP P33173
K	-2	PRO	-	see remark 999	UNP P33173
K	-1	ILE	-	see remark 999	UNP P33173

Continued on next page...

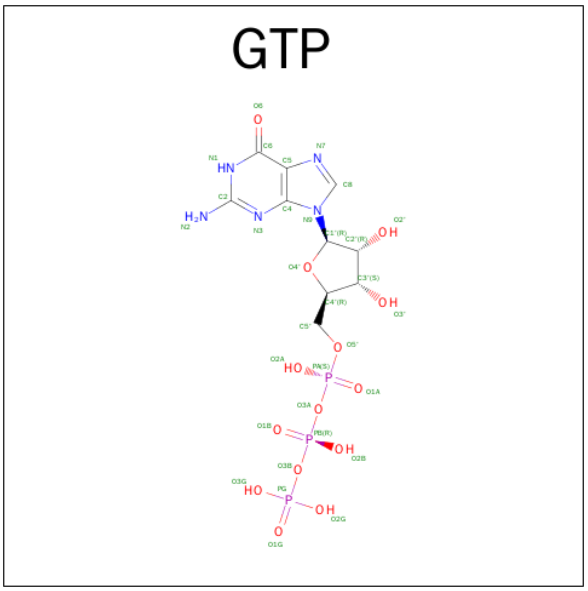
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
K	0	ASN	-	see remark 999	UNP P33173
K	1	MET	-	see remark 999	UNP P33173
K	2	PRO	-	see remark 999	UNP P33173
K	202	ALA	PRO	engineered	UNP P33173
K	356	ASN	-	SEE REMARK 999	UNP P33173
K	357	THR	-	SEE REMARK 999	UNP P33173
K	358	VAL	-	SEE REMARK 999	UNP P33173
K	359	SER	-	SEE REMARK 999	UNP P33173
K	360	VAL	-	SEE REMARK 999	UNP P33173
K	361	ASN	-	SEE REMARK 999	UNP P33173
K	362	LEU	-	SEE REMARK 999	UNP P33173
K	363	GLU	-	SEE REMARK 999	UNP P33173
K	364	LEU	-	SEE REMARK 999	UNP P33173
K	365	THR	-	SEE REMARK 999	UNP P33173
K	366	ALA	-	SEE REMARK 999	UNP P33173
K	367	GLU	-	SEE REMARK 999	UNP P33173
K	368	GLU	-	SEE REMARK 999	UNP P33173
K	369	TRP	-	SEE REMARK 999	UNP P33173
K	370	LYS	-	SEE REMARK 999	UNP P33173
K	371	LYS	-	SEE REMARK 999	UNP P33173
K	372	LYS	-	SEE REMARK 999	UNP P33173
K	373	HIS	-	SEE REMARK 999	UNP P33173
K	374	HIS	-	SEE REMARK 999	UNP P33173
K	375	HIS	-	SEE REMARK 999	UNP P33173
K	376	HIS	-	SEE REMARK 999	UNP P33173
K	377	HIS	-	SEE REMARK 999	UNP P33173
K	378	HIS	-	SEE REMARK 999	UNP P33173

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

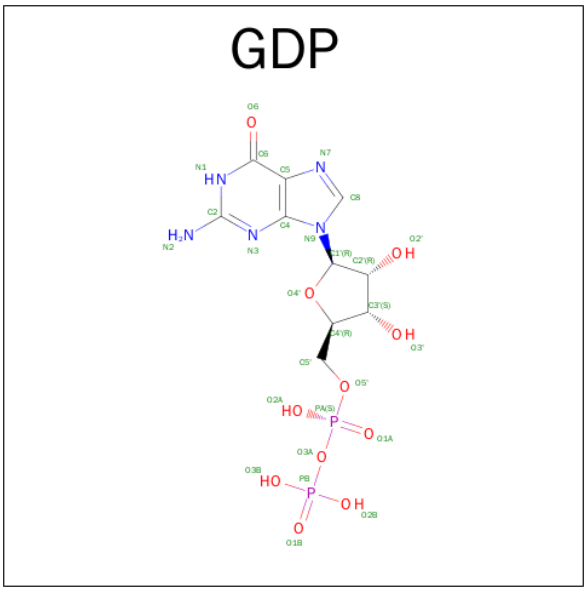
Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
4	K	1	1	1	0

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



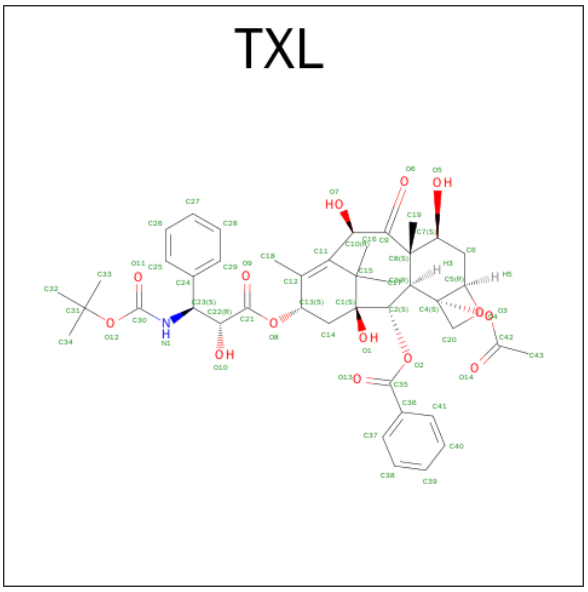
Mol	Chain	Residues	Atoms					AltConf
5	A	1	Total	C	N	O	P	0
			32	10	5	14	3	

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



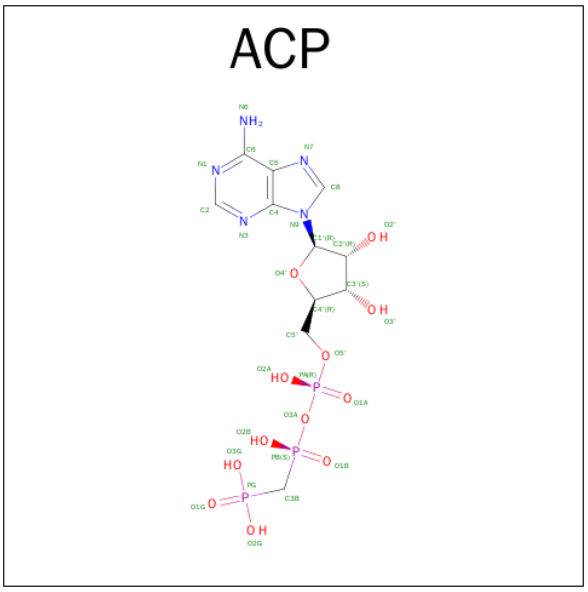
Mol	Chain	Residues	Atoms					AltConf
6	B	1	Total	C	N	O	P	0
			28	10	5	11	2	

- Molecule 7 is TAXOTERE (three-letter code: TXL) (formula: $C_{43}H_{53}NO_{14}$).



Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
7	B	1	58	43	1	14	0

- Molecule 8 is PHOSPHOMETHYLPHOSPHONIC ACID ADENYLATE ESTER (three-letter code: ACP) (formula: $C_{11}H_{18}N_5O_{12}P_3$).



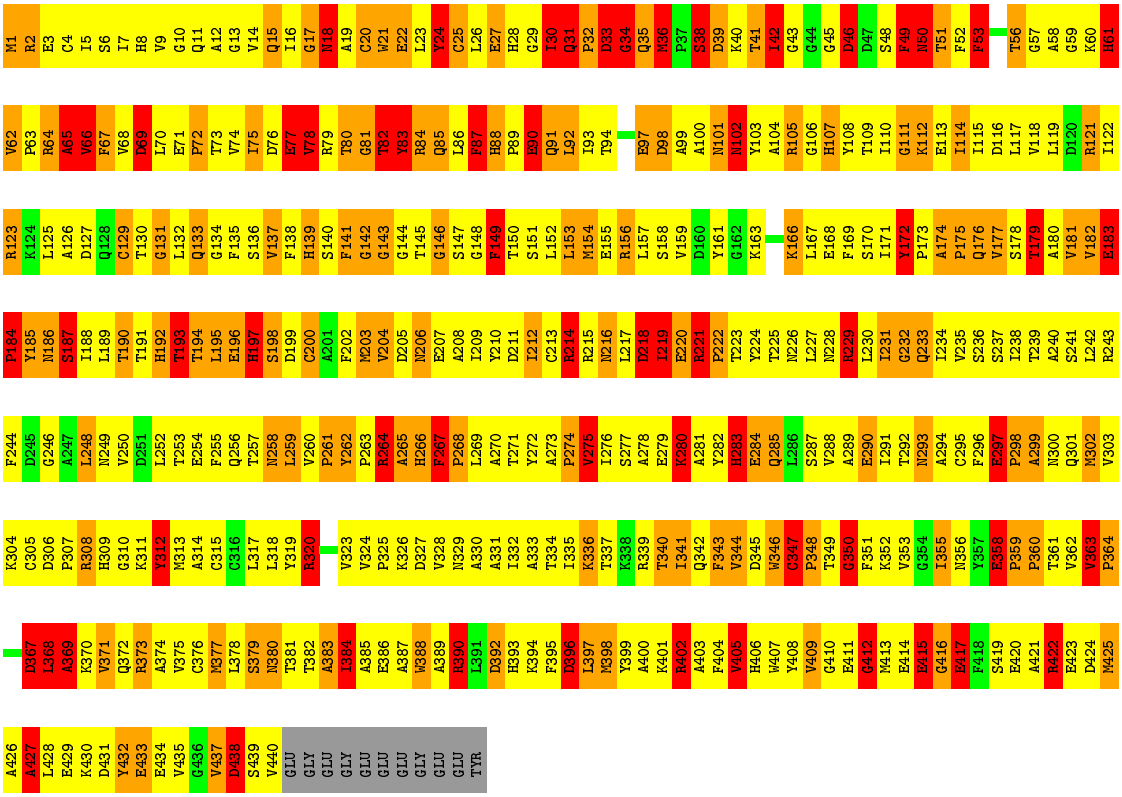
Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
8	K	1	31	11	5	12	3	0

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 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.

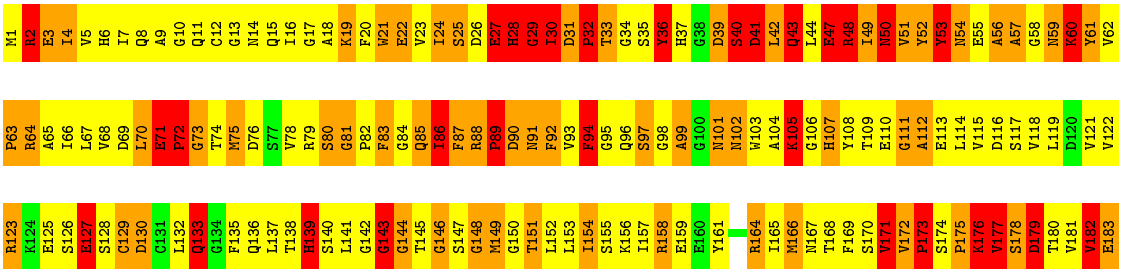
- Molecule 1: TUBULIN ALPHA CHAIN

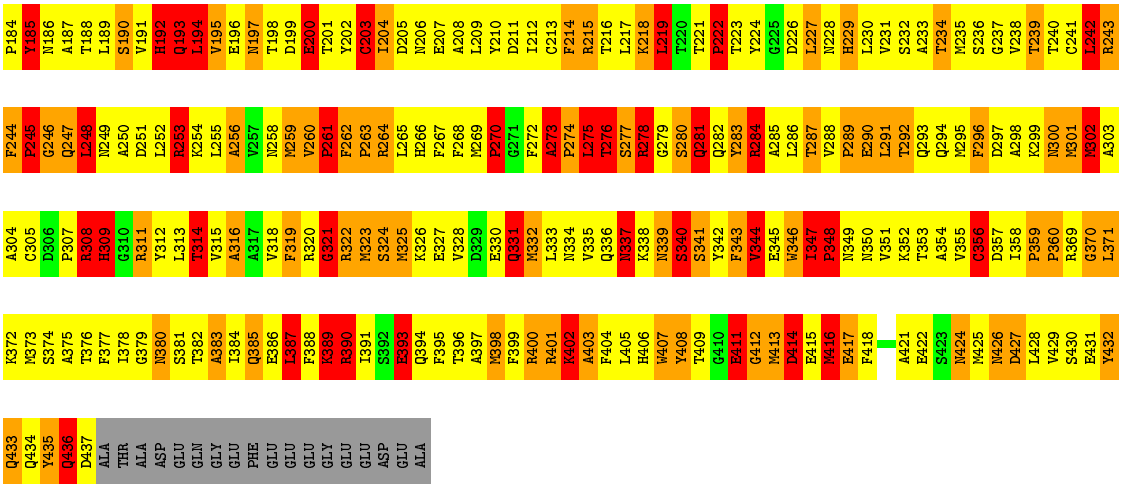
Chain A: 7% 51% 26% 14%



- Molecule 2: TUBULIN BETA CHAIN

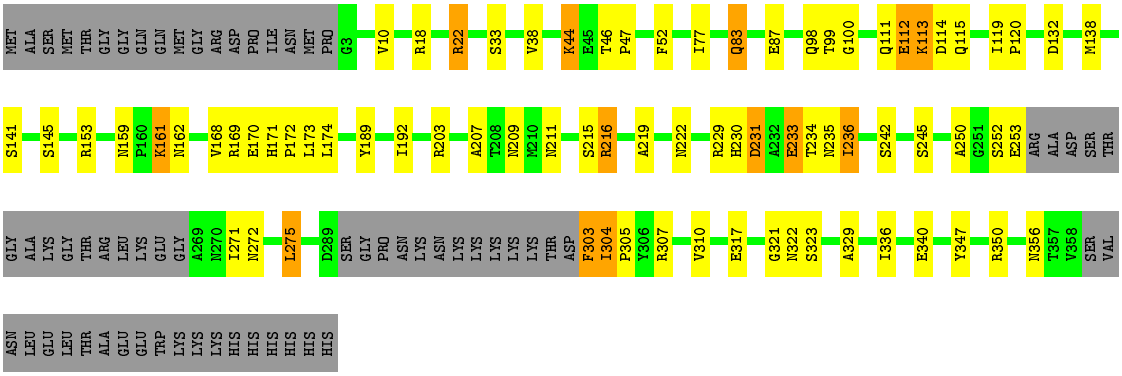
Chain B: 5% 47% 28% 16%





- Molecule 3: KINESIN-LIKE PROTEIN KIF1A

Chain K:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of images	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	Not provided	Depositor
Microscope	JEOL JEM-2010F	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	1000	Depositor
Minimum defocus (nm)	1700	Depositor
Maximum defocus (nm)	3000	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: GDP, GTP, MG, TXL, ACP

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	2.34	109/3508 (3.1%)	2.76	211/4762 (4.4%)
2	B	2.47	110/3434 (3.2%)	3.07	266/4652 (5.7%)
3	K	0.47	0/2708	0.73	2/3655 (0.1%)
All	All	2.05	219/9650 (2.3%)	2.51	479/13069 (3.7%)

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	57
2	B	0	59
All	All	0	116

The worst 5 of 219 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	278	ARG	CA-CB	34.91	2.30	1.53
2	B	105	LYS	C-N	-29.41	0.80	1.33
2	B	73	GLY	C-N	-28.02	0.69	1.34
1	A	38	SER	C-N	-27.48	0.70	1.34
1	A	347	CYS	C-N	-23.39	0.89	1.34

The worst 5 of 479 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	368	LEU	O-C-N	-54.19	35.99	122.70
1	A	363	VAL	C-N-CD	-48.68	13.50	120.60
2	B	273	ALA	C-N-CD	-46.49	18.32	120.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	105	LYS	O-C-N	-44.51	47.53	123.20
2	B	88	ARG	C-N-CD	-44.07	23.64	120.60

There are no chirality outliers.

5 of 116 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	18	ASN	Mainchain
1	A	24	TYR	Mainchain,Peptide
1	A	30	ILE	Mainchain,Peptide
1	A	34	GLY	Peptide
1	A	36	MET	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	3430	0	3252	1687	0
2	B	3359	0	3171	1861	0
3	K	2667	0	2614	154	0
4	K	1	0	0	0	0
5	A	32	0	11	21	0
6	B	28	0	12	17	0
7	B	58	0	51	57	0
8	K	31	0	14	2	0
All	All	9606	0	9125	3576	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 191.

The worst 5 of 3576 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)
2:B:405:LEU:HD22	2:B:418:PHE:CZ	1.17	1.68
2:B:151:THR:CB	2:B:192:HIS:CD2	1.75	1.68

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:229:HIS:CE1	7:B:502:TXL:H343	1.28	1.67
1:A:115:ILE:HD12	1:A:152:LEU:CG	1.26	1.64
2:B:184:PRO:HG2	2:B:399:PHE:CE2	1.12	1.63

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	438/451 (97%)	322 (74%)	63 (14%)	53 (12%)	0 8
2	B	425/445 (96%)	298 (70%)	48 (11%)	79 (19%)	0 3
3	K	332/394 (84%)	319 (96%)	11 (3%)	2 (1%)	30 74
All	All	1195/1290 (93%)	939 (79%)	122 (10%)	134 (11%)	1 11

5 of 134 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	31	GLN
1	A	33	ASP
1	A	35	GLN
1	A	39	ASP
1	A	42	ILE

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	369/377 (98%)	351 (95%)	18 (5%)	31 67
2	B	368/381 (97%)	348 (95%)	20 (5%)	27 64
3	K	300/345 (87%)	273 (91%)	27 (9%)	12 44
All	All	1037/1103 (94%)	972 (94%)	65 (6%)	29 59

5 of 65 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	302	MET
2	B	432	TYR
3	K	271	ILE
2	B	308	ARG
2	B	416	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	28	HIS
2	B	102	ASN
3	K	159	ASN
2	B	59	ASN
1	A	206	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

Of 5 ligands modelled in this entry, 1 is 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	$\#\|Z\| > 2$
5	GTP	A	500	1	26,34,34	1.74	6 (23%)	29,54,54	2.63	3 (10%)
6	GDP	B	501	2	24,30,30	1.84	6 (25%)	26,47,47	2.75	6 (23%)
7	TXL	B	502	2	61,63,63	4.00	45 (73%)	100,100,100	2.93	47 (47%)
8	ACP	K	503	4	29,33,33	1.43	5 (17%)	29,52,52	2.43	1 (3%)

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
5	GTP	A	500	1	-	0/18/38/38	0/3/3/3
6	GDP	B	501	2	-	0/12/32/32	0/3/3/3
7	TXL	B	502	2	-	0/38/124/124	0/4/6/6
8	ACP	K	503	4	-	0/15/38/38	0/3/3/3

The worst 5 of 62 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	B	502	TXL	O3-C4	-8.79	1.26	1.46
7	B	502	TXL	C37-C36	-8.02	1.26	1.39
7	B	502	TXL	C25-C24	-7.32	1.27	1.39
7	B	502	TXL	O5-C7	-6.60	1.32	1.43
7	B	502	TXL	C13-C12	-6.56	1.37	1.51

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	K	503	ACP	N3-C2-N1	-12.22	119.27	128.87
5	A	500	GTP	C5-C6-N1	-9.86	110.63	123.52
6	B	501	GDP	C5-C6-N1	-9.42	111.20	123.52
7	B	502	TXL	C38-C37-C36	-8.03	110.78	120.35
7	B	502	TXL	C13-C12-C11	-6.04	108.09	117.72

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 97 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	500	GTP	21	0
6	B	501	GDP	17	0
7	B	502	TXL	57	0
8	K	503	ACP	2	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.