



wwPDB/EMDataBank EM Map/Model Validation Summary Report ⓘ

Nov 9, 2017 – 08:19 AM EST

PDB ID : 5OGC
EMDB ID: : EMD-3803
Title : Molecular basis of human kinesin-8 function and inhibition
Authors : Locke, J.; Joseph, A.P.; Topf, M.; Moores, C.A.
Deposited on : unknown
Resolution : 4.80 Å(reported)

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
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20030345

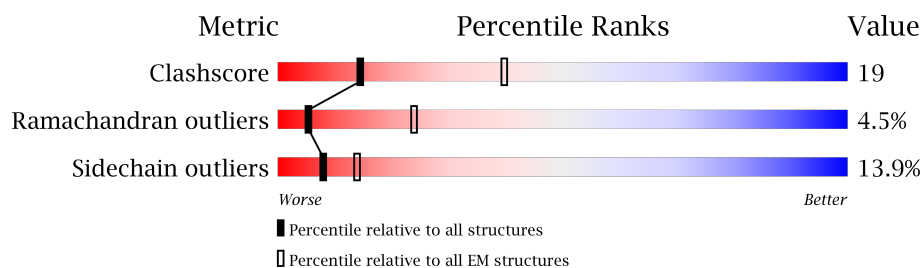
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.80 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	125131	1336
Ramachandran outliers	121729	1120
Sidechain outliers	121581	1026

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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	1-K	377	41% 34% 10% • 13%
1	2-K	377	55% 28% • 13%
2	1-A	451	50% 32% 7% • 9%
2	2-A	451	49% 33% 8% • 9%
3	1-B	445	52% 34% 9% • •
3	2-B	445	52% 34% 9% • •

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
7	GTP	1-A	503	-	-	X	-
7	GTP	2-A	503	-	-	X	-

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 18580 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 Kinesin-like protein KIF18A.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	1-K	328	Total	C	N	O	S	0	0
			2569	1605	463	489	12		
1	2-K	328	Total	C	N	O	S	0	0
			2569	1605	463	489	12		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
K	-2	GLY	-	expression tag	UNP Q8NI77
K	-1	SER	-	expression tag	UNP Q8NI77
K	0	HIS	-	expression tag	UNP Q8NI77

- Molecule 2 is a protein called Tubulin alpha chain.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	1-A	412	Total	C	N	O	S	0	0
			3227	2043	551	613	20		
2	2-A	412	Total	C	N	O	S	0	0
			3227	2043	551	613	20		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	136	SER	LEU	conflict	UNP F2Z4C1
A	265	GLY	ILE	conflict	UNP F2Z4C1
A	358	GLU	GLN	conflict	UNP F2Z4C1

- Molecule 3 is a protein called Tubulin beta chain.

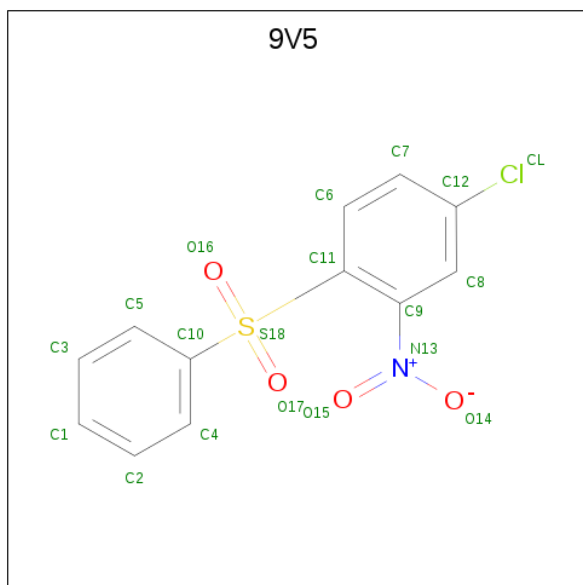
Mol	Chain	Residues	Atoms					AltConf	Trace
3	1-B	426	Total	C	N	O	S	0	0
			3351	2105	575	646	25		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
3	2-B	426	Total	C	N	O	S	0	0
			3351	2105	575	646	25		

- Molecule 4 is 4-chloranyl-2-nitro-1-(phenylsulfonyl)benzene (three-letter code: 9V5) (formula: C₁₂H₈ClNO₄S).



Mol	Chain	Residues	Atoms					AltConf	
4	1-K	1	Total	C	Cl	N	O	S	0
			19	12	1	1	4	1	
4	2-K	1	Total	C	Cl	N	O	S	0
			19	12	1	1	4	1	

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
5	2-A	1	Total	Zn	0
			1	1	
5	1-A	1	Total	Zn	0
			1	1	

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

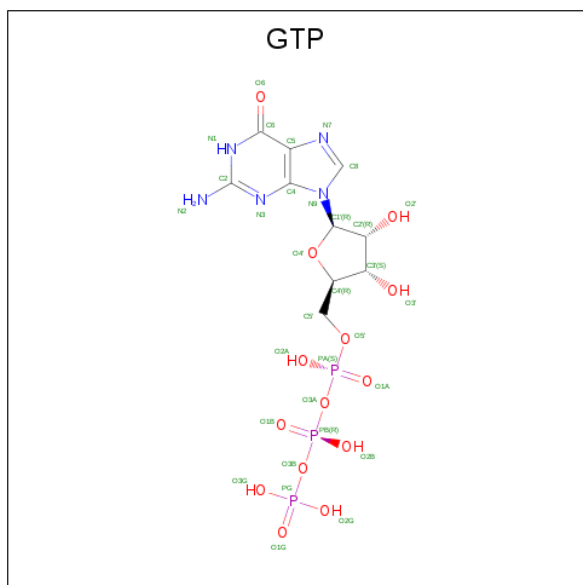
Mol	Chain	Residues	Atoms		AltConf
6	2-A	1	Total	Mg	0
			1	1	

Continued on next page...

Continued from previous page...

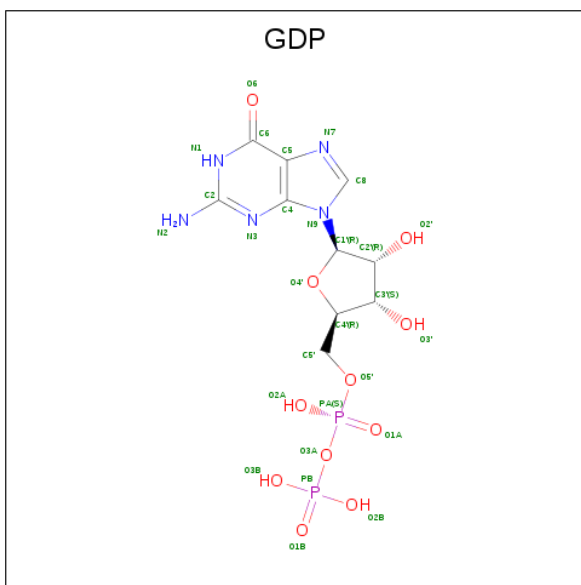
Mol	Chain	Residues	Atoms		AltConf
6	1-A	1	Total	Mg	0
			1	1	

- Molecule 7 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula: $C_{10}H_{16}N_5O_{14}P_3$).



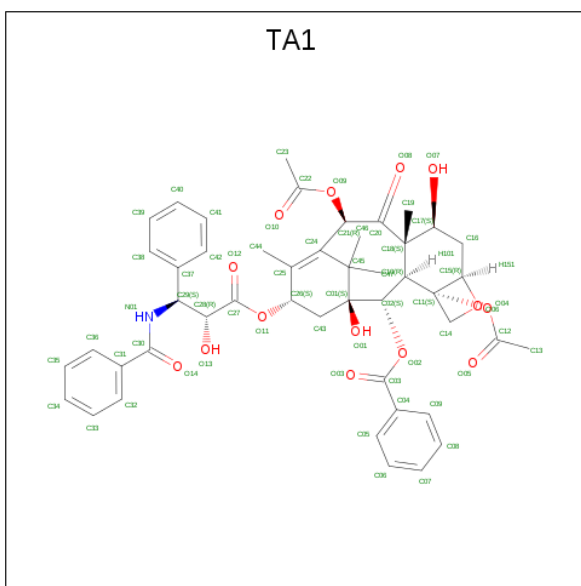
Mol	Chain	Residues	Atoms					AltConf
7	1-A	1	Total	C	N	O	P	0
			32	10	5	14	3	
7	2-A	1	Total	C	N	O	P	0
			32	10	5	14	3	

- Molecule 8 is GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: $C_{10}H_{15}N_5O_{11}P_2$).



Mol	Chain	Residues	Atoms					AltConf
8	1-B	1	Total 28	C 10	N 5	O 11	P 2	0
8	2-B	1	Total 28	C 10	N 5	O 11	P 2	0

- Molecule 9 is TAXOL (three-letter code: TA1) (formula: $C_{47}H_{51}NO_{14}$).

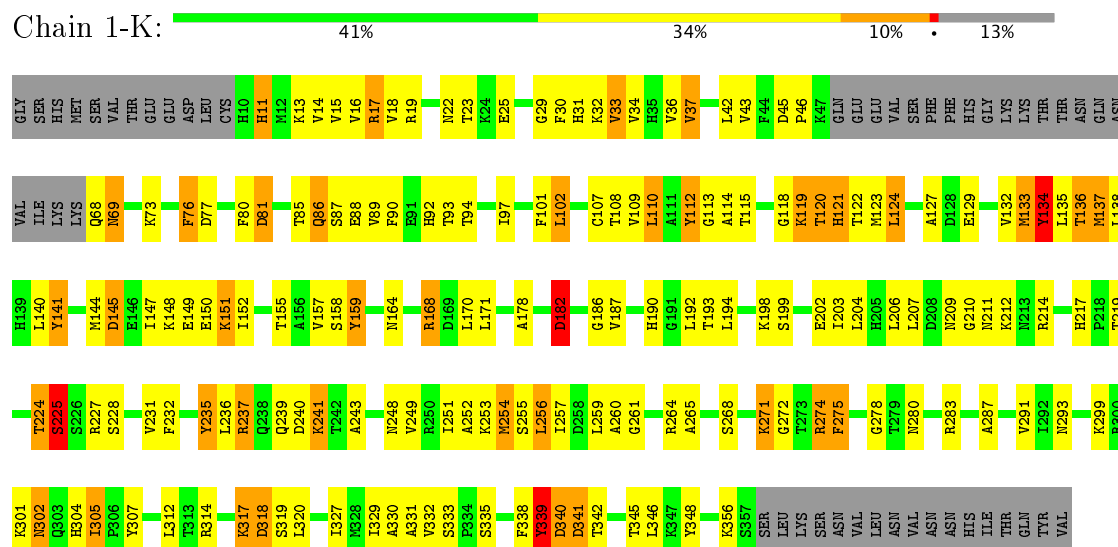


Mol	Chain	Residues	Atoms				AltConf
9	1-B	1	Total 62	C 47	N 1	O 14	0
9	2-B	1	Total 62	C 47	N 1	O 14	0

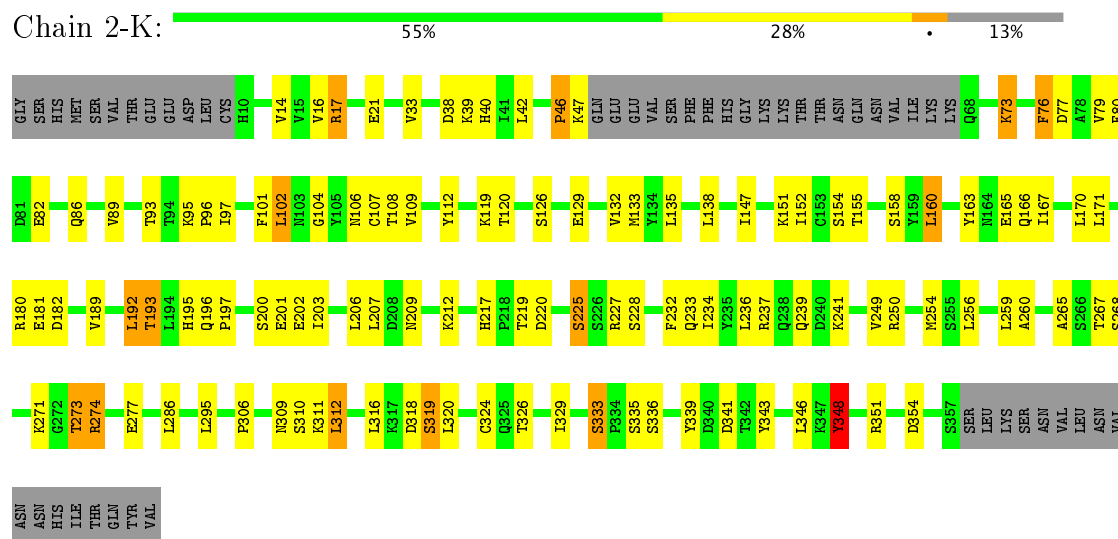
3 Residue-property plots

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: Kinesin-like protein KIF18A

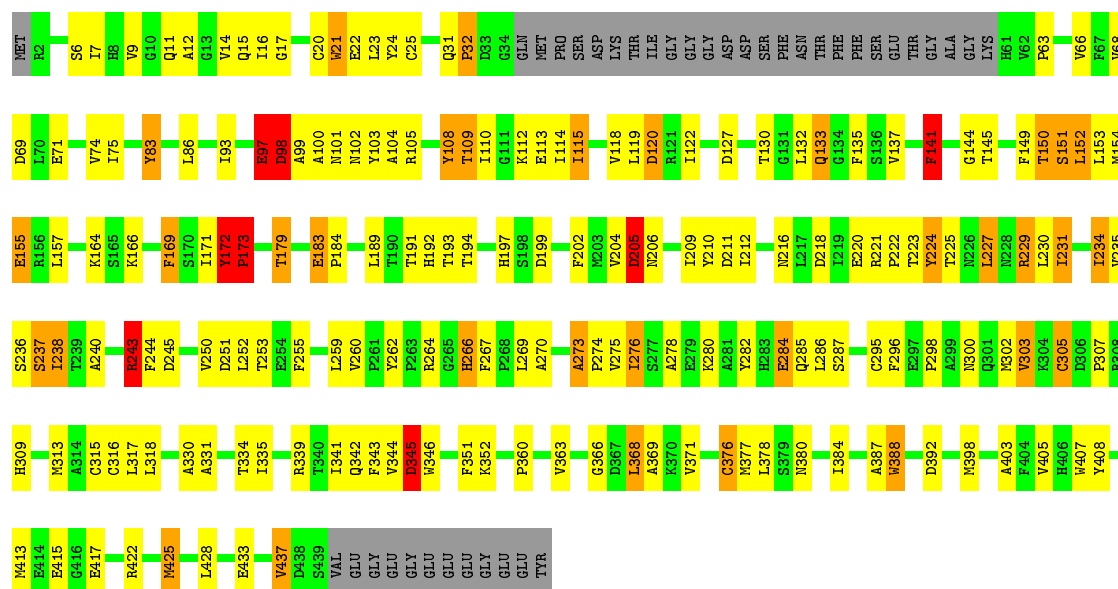


• Molecule 1: Kinesin-like protein KIF18A



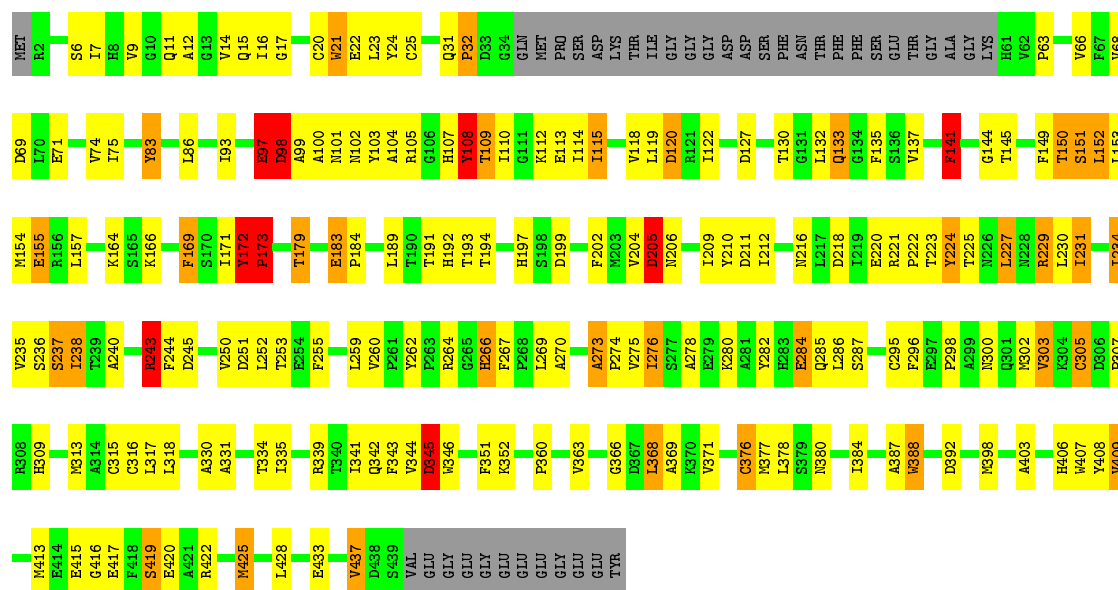
• Molecule 2: Tubulin alpha chain





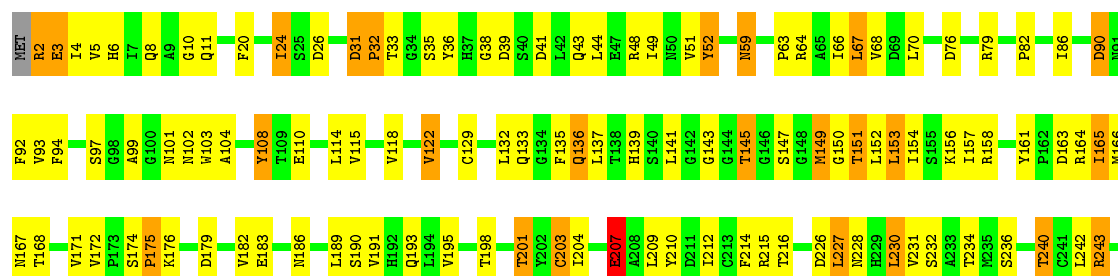
• Molecule 2: Tubulin alpha chain

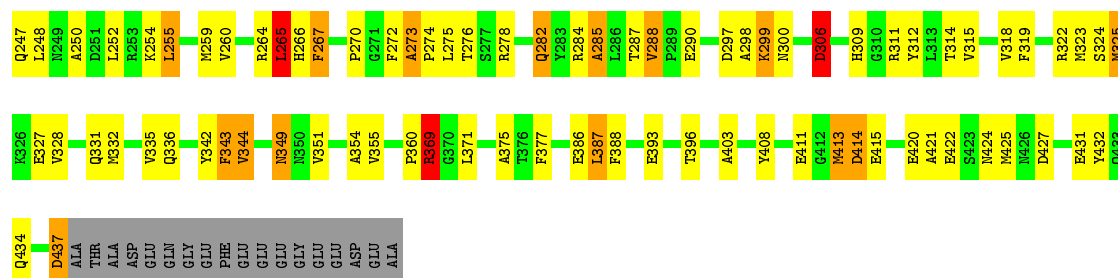
Chain 2-A: 49% 33% 8% 9%



• Molecule 3: Tubulin beta chain

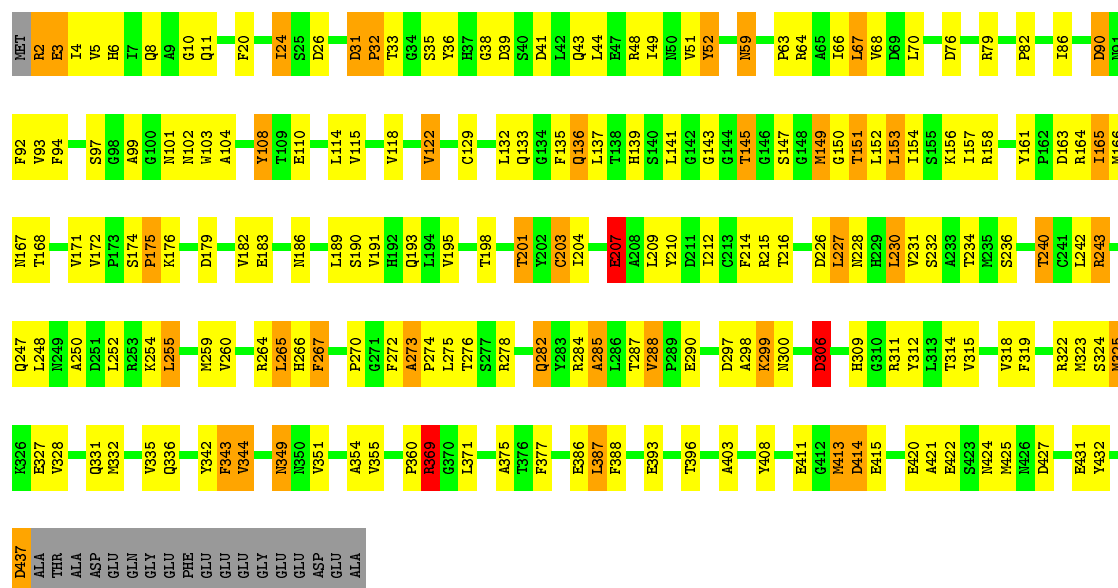
Chain 1-B: 52% 34% 9%





• Molecule 3: Tubulin beta chain

Chain 2-B: 52% 34% 9%



4 Experimental information

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	HELICAL, twist=0°, rise=81 Å, axial sym=C1	Depositor
Number of segments used	135382	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 ($e^-/\text{\AA}^2$)	5	Depositor
Minimum defocus (nm)	500	Depositor
Maximum defocus (nm)	3500	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, 9V5, ZN, TA1, MG, 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	1-K	0.65	0/2611	1.14	8/3524 (0.2%)
1	2-K	0.32	0/2611	0.49	1/3524 (0.0%)
2	1-A	0.70	0/3300	1.19	10/4482 (0.2%)
2	2-A	0.71	0/3300	1.20	11/4482 (0.2%)
3	1-B	0.74	2/3426 (0.1%)	1.17	8/4642 (0.2%)
3	2-B	0.74	2/3426 (0.1%)	1.17	9/4642 (0.2%)
All	All	0.67	4/18674 (0.0%)	1.10	47/25296 (0.2%)

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	1-K	0	27
1	2-K	1	3
2	1-A	0	27
2	2-A	0	32
3	1-B	0	28
3	2-B	0	28
All	All	1	145

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	1-B	431	GLU	CD-OE2	-9.96	1.14	1.25
3	2-B	431	GLU	CD-OE2	-9.96	1.14	1.25
3	1-B	431	GLU	CD-OE1	9.24	1.35	1.25
3	2-B	431	GLU	CD-OE1	9.24	1.35	1.25

The worst 5 of 47 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	1-B	64	ARG	NE-CZ-NH1	11.25	125.93	120.30
3	2-B	64	ARG	NE-CZ-NH1	11.25	125.93	120.30
2	1-A	243	ARG	NE-CZ-NH1	9.74	125.17	120.30
2	2-A	243	ARG	NE-CZ-NH1	9.74	125.17	120.30
1	1-K	340	ASP	CB-CG-OD1	-8.19	110.93	118.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	2-K	47	LYS	CA

5 of 145 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	1-K	120	THR	Mainchain
1	1-K	69	ASN	Sidechain
1	1-K	77	ASP	Peptide
1	1-K	81	ASP	Sidechain
1	1-K	86	GLN	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	1-K	2569	0	2587	152	0
1	2-K	2569	0	2587	65	0
2	1-A	3227	0	3143	126	0
2	2-A	3227	0	3143	133	0
3	1-B	3351	0	3229	125	0
3	2-B	3351	0	3229	122	0
4	1-K	19	0	0	0	0
4	2-K	19	0	0	5	0
5	1-A	1	0	0	0	0
5	2-A	1	0	0	0	0
6	1-A	1	0	0	0	0
6	2-A	1	0	0	0	0
7	1-A	32	0	12	10	0
7	2-A	32	0	12	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	1-B	28	0	12	2	0
8	2-B	28	0	12	2	0
9	1-B	62	0	51	6	0
9	2-B	62	0	51	6	0
All	All	18580	0	18068	707	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 19.

The worst 5 of 707 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:K:33:VAL:HG13	1:K:34:VAL:H	1.05	1.12
3:B:172:VAL:HG11	3:B:387:LEU:HD21	1.32	1.11
3:B:172:VAL:HG11	3:B:387:LEU:HD21	1.32	1.11
1:K:31:HIS:HE1	1:K:46:PRO:O	1.40	1.02
1:K:291:VAL:HG13	1:K:305:ILE:HD11	1.41	0.97

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	1-K	324/377 (86%)	288 (89%)	28 (9%)	8 (2%)	6	42
1	2-K	324/377 (86%)	292 (90%)	28 (9%)	4 (1%)	15	58
2	1-A	408/451 (90%)	316 (78%)	66 (16%)	26 (6%)	1	23
2	2-A	408/451 (90%)	317 (78%)	65 (16%)	26 (6%)	1	23
3	1-B	424/445 (95%)	338 (80%)	66 (16%)	20 (5%)	3	29
3	2-B	424/445 (95%)	338 (80%)	66 (16%)	20 (5%)	3	29

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	2312/2546 (91%)	1889 (82%)	319 (14%)	104 (4%)	5	30

5 of 104 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	1-K	33	VAL
1	1-K	268	SER
2	1-A	97	GLU
2	1-A	108	TYR
2	1-A	109	THR

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	1-K	287/334 (86%)	251 (88%)	36 (12%)	5	28
1	2-K	287/334 (86%)	243 (85%)	44 (15%)	3	21
2	1-A	347/377 (92%)	303 (87%)	44 (13%)	5	27
2	2-A	347/377 (92%)	303 (87%)	44 (13%)	5	27
3	1-B	367/381 (96%)	312 (85%)	55 (15%)	3	22
3	2-B	367/381 (96%)	312 (85%)	55 (15%)	3	22
All	All	2002/2184 (92%)	1724 (86%)	278 (14%)	8	25

5 of 278 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	1-B	343	PHE
1	2-K	192	LEU
3	2-B	240	THR
3	1-B	387	LEU
1	2-K	102	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 50 such sidechains are listed below:

Mol	Chain	Res	Type
3	1-B	385	GLN
1	2-K	293	ASN
3	2-B	385	GLN
3	1-B	424	ASN
3	1-B	434	GLN

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 12 ligands modelled in this entry, 4 are monoatomic - leaving 8 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
7	GTP	1-A	503	6	27,34,34	1.53	2 (7%)	27,54,54	2.16	5 (18%)
8	GDP	1-B	600	-	25,30,30	2.61	8 (32%)	26,47,47	3.64	9 (34%)
9	TA1	1-B	601	-	68,68,68	1.97	20 (29%)	105,105,105	1.33	9 (8%)
4	9V5	1-K	401	-	18,20,20	0.63	0	24,29,29	1.48	3 (12%)
7	GTP	2-A	503	6	27,34,34	1.53	2 (7%)	27,54,54	2.16	5 (18%)
8	GDP	2-B	600	-	25,30,30	2.61	8 (32%)	26,47,47	3.64	9 (34%)
9	TA1	2-B	601	-	68,68,68	1.97	20 (29%)	105,105,105	1.33	9 (8%)
4	9V5	2-K	401	-	18,20,20	2.86	4 (22%)	24,29,29	1.90	4 (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
7	GTP	1-A	503	6	-	0/18/38/38	0/3/3/3
8	GDP	1-B	600	-	-	0/12/32/32	0/3/3/3
9	TA1	1-B	601	-	-	0/41/127/127	0/5/7/7
4	9V5	1-K	401	-	-	1/14/16/16	0/2/2/2
7	GTP	2-A	503	6	-	0/18/38/38	0/3/3/3
8	GDP	2-B	600	-	-	0/12/32/32	0/3/3/3
9	TA1	2-B	601	-	-	0/41/127/127	0/5/7/7
4	9V5	2-K	401	-	-	0/14/16/16	0/2/2/2

The worst 5 of 64 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	2-B	601	TA1	C08-C07	-5.12	1.25	1.38
9	1-B	601	TA1	C08-C07	-5.12	1.25	1.38
8	2-B	600	GDP	PB-O2B	-3.54	1.40	1.54
8	1-B	600	GDP	PB-O2B	-3.54	1.40	1.54
9	2-B	601	TA1	C04-C03	-2.31	1.44	1.49

The worst 5 of 53 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	2-B	600	GDP	C6-C5-C4	-11.37	109.54	120.84
8	1-B	600	GDP	C6-C5-C4	-11.37	109.54	120.84
7	2-A	503	GTP	C5-C6-N1	-7.10	113.37	123.48
7	1-A	503	GTP	C5-C6-N1	-7.10	113.37	123.48
8	2-B	600	GDP	N2-C2-N1	-5.89	107.82	117.24

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	1-K	401	9V5	C9-C11-S18-C10

There are no ring outliers.

7 monomers are involved in 41 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	1-A	503	GTP	10	0
8	1-B	600	GDP	2	0
9	1-B	601	TA1	6	0
7	2-A	503	GTP	10	0
8	2-B	600	GDP	2	0
9	2-B	601	TA1	6	0
4	2-K	401	9V5	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.