wwPDB EM Map/Model Validation Report (i)

Apr 10, 2016 – 01:43 PM BST

PDB ID	:	3IZ0
EMDB ID:	:	EMD-5223
Title	:	Human Ndc80 Bonsai Decorated Microtubule
Authors	:	Alushin, G.M.; Ramey, V.H.; Pasqualato, S.; Ball, D.A.; Grigorieff, N.; Musacchio, A.; Nogales, E.
Deposited on	:	2010-08-09
Resolution	:	8.60 Å(reported)
Based on PDB ID	:	2VE7, 1JFF

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

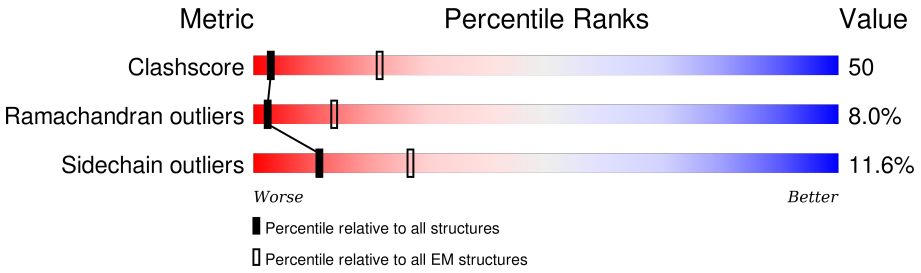
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk27241

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 8.60 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	16% 57% 17% • 9%
2	B	445	20% 56% 18% • •
3	C	315	48% 9% • 42%
3	E	315	49% 8% • 42%
4	D	250	44% 14% • 39%
4	F	250	43% 15% • 39%

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 12154 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 alpha tubulin, Chain A from PDB 1JFF.

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 beta tubulin, Chain B from PDB 1JFF.

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 NDC80-SPC25 chimera protein, Chain B from PDB 2VE7 (Ndc80 bonsai).

Mol	Chain	Residues	Atoms					AltConf	Trace
3	C	182	Total	C	N	O	S	0	0
			1465	950	229	274	12		
3	E	182	Total	C	N	O	S	0	0
			1465	950	229	274	12		

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	79	MET	-	EXPRESSION TAG	UNP Q05DQ6
C	393	GLN	ASN	CONFLICT	UNP Q9HBM1
E	79	MET	-	EXPRESSION TAG	UNP Q05DQ6
E	393	GLN	ASN	CONFLICT	UNP Q9HBM1

- Molecule 4 is a protein called NUF2-SPC24 chimera protein, Chain D from PDB 2VE7 (Ndc80 bonsai).

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	153	Total	C	N	O	S	0	0
			1261	818	215	217	11		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
4	F	153	Total	C	N	O	S	0	0
			1261	818	215	217	11		

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	-4	GLY	-	EXPRESSION TAG	UNP B1AQT4
D	-3	PRO	-	EXPRESSION TAG	UNP B1AQT4
D	-2	LEU	-	EXPRESSION TAG	UNP B1AQT4
D	-1	GLY	-	EXPRESSION TAG	UNP B1AQT4
D	0	SER	-	EXPRESSION TAG	UNP B1AQT4
D	72	GLY	GLU	ENGINEERED MUTATION	UNP B1AQT4
D	200	GLU	ASP	CONFLICT	UNP C9JGC4
F	-4	GLY	-	EXPRESSION TAG	UNP B1AQT4
F	-3	PRO	-	EXPRESSION TAG	UNP B1AQT4
F	-2	LEU	-	EXPRESSION TAG	UNP B1AQT4
F	-1	GLY	-	EXPRESSION TAG	UNP B1AQT4
F	0	SER	-	EXPRESSION TAG	UNP B1AQT4
F	72	GLY	GLU	ENGINEERED MUTATION	UNP B1AQT4
F	200	GLU	ASP	CONFLICT	UNP C9JGC4

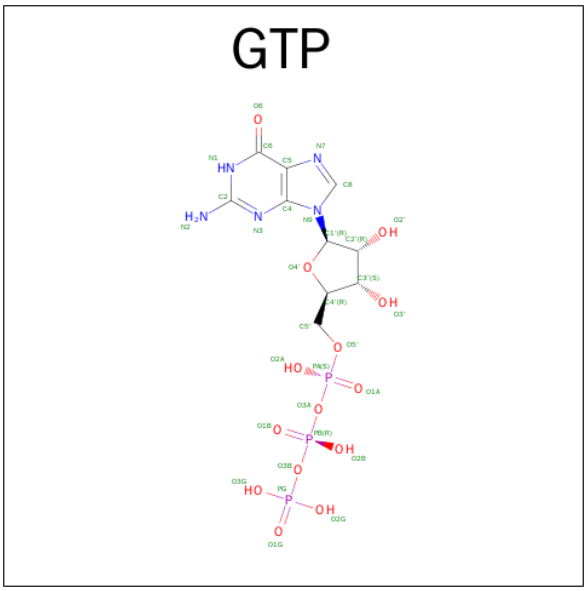
- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

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

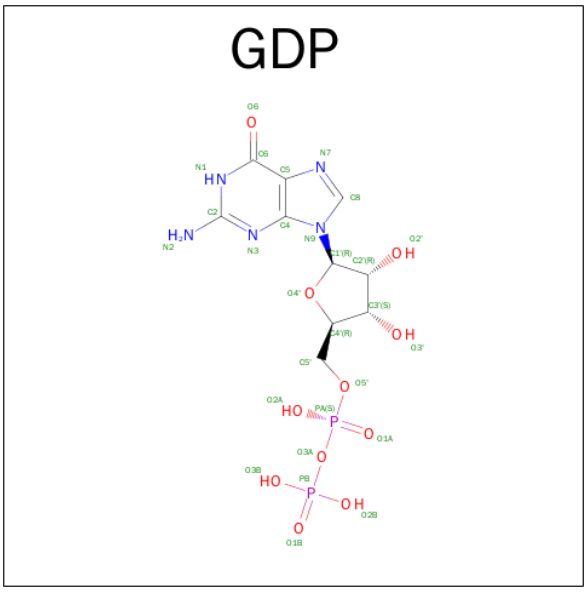
Mol	Chain	Residues	Atoms		AltConf
6	A	1	Total	Mg	0
			1	1	

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



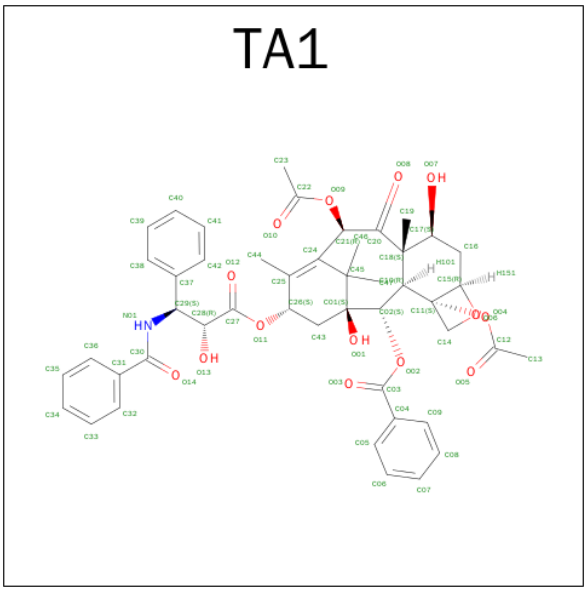
Mol	Chain	Residues	Atoms					AltConf
7	A	1	Total	C	N	O	P	0
			32	10	5	14	3	

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

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



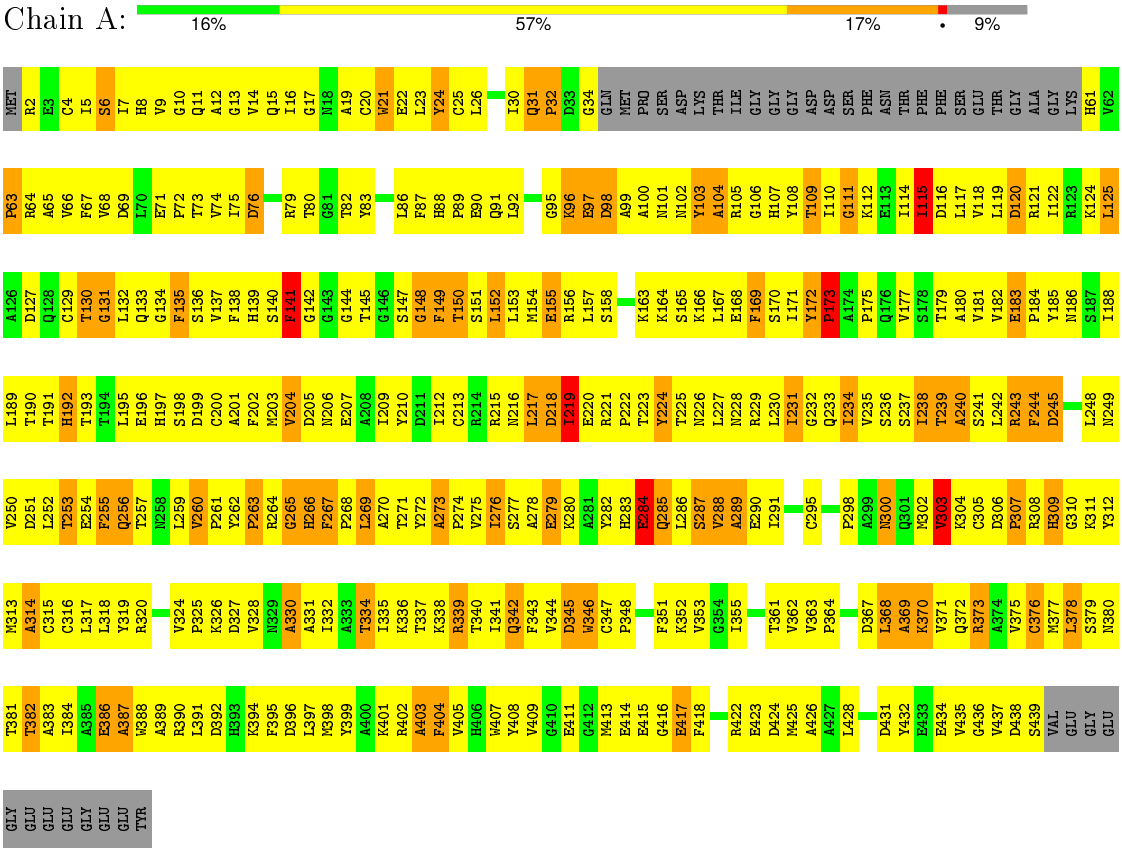
Mol	Chain	Residues	Atoms				AltConf
			Total	C	N	O	
9	B	1	62	47	1	14	0

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.

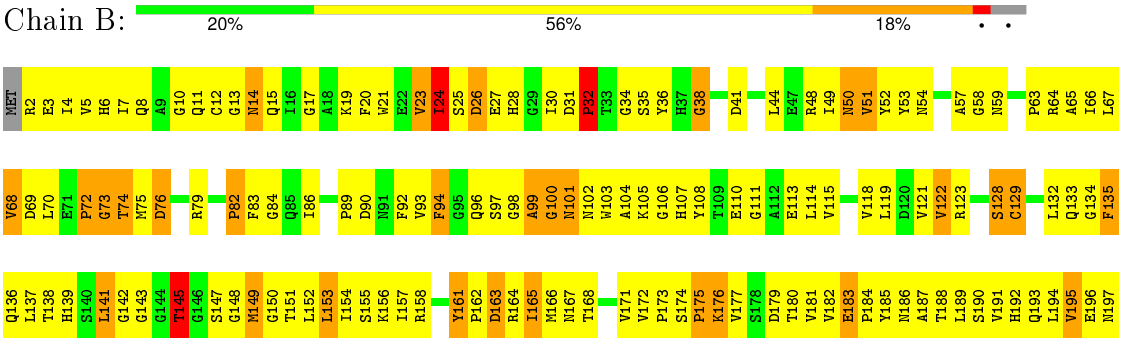
- Molecule 1: alpha tubulin, Chain A from PDB 1JFF

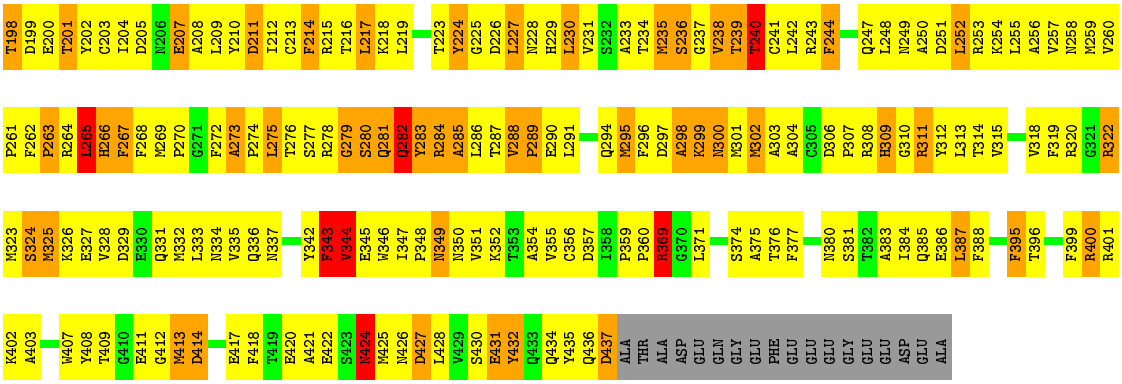
Chain A:



- Molecule 2: beta tubulin, Chain B from PDB 1JFF

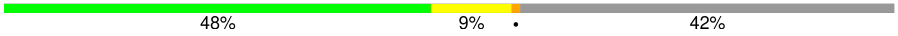
Chain B:

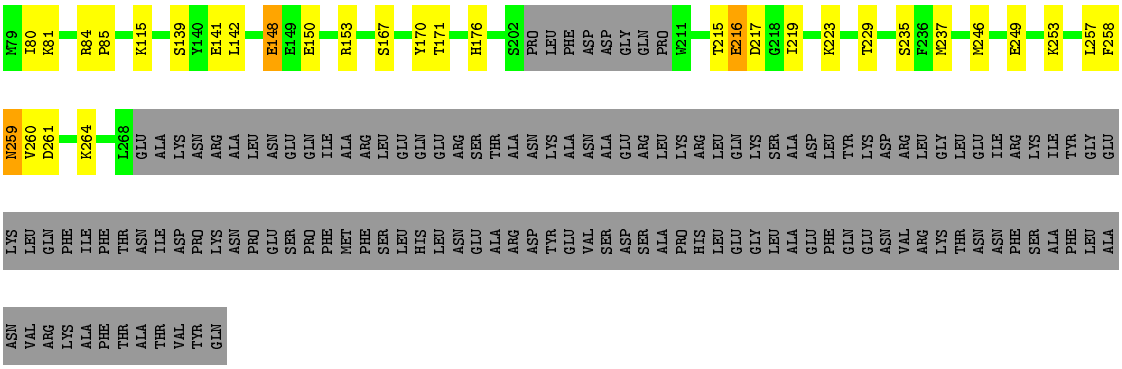




• Molecule 3: NDC80-SPC25 chimera protein, Chain B from PDB 2VE7 (Ndc80 bonsai)

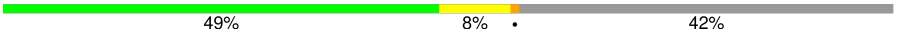
Chain C:

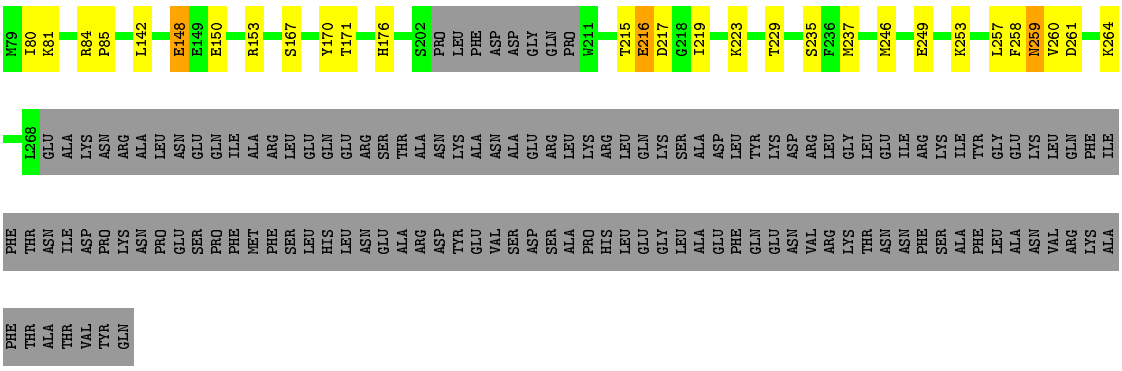




• Molecule 3: NDC80-SPC25 chimera protein, Chain B from PDB 2VE7 (Ndc80 bonsai)

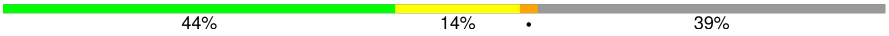
Chain E:

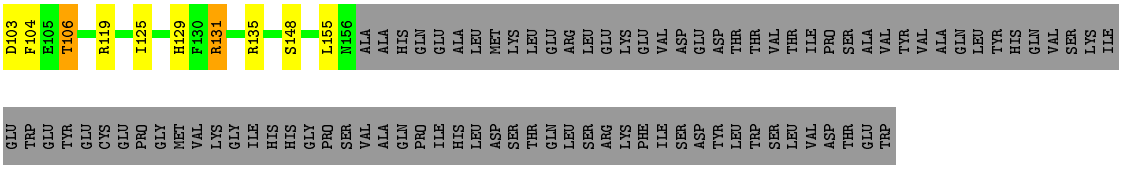




• Molecule 4: NUF2-SPC24 chimera protein, Chain D from PDB 2VE7 (Ndc80 bonsai)

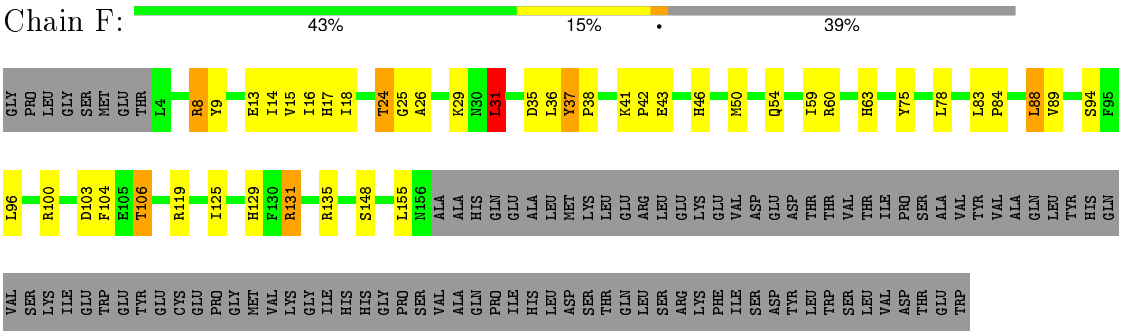
Chain D:





• Molecule 4: NUF2-SPC24 chimera protein, Chain D from PDB 2VE7 (Ndc80 bonsai)

Chain F:



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	FSC at 0.143 cutoff	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TECNAI F20	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	15	Depositor
Minimum defocus (nm)	1200	Depositor
Maximum defocus (nm)	2200	Depositor
Magnification	50000	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, ZN, TA1, 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	#$\|Z\|>2$	RMSZ	#$\|Z\|>2$
1	A	0.50	0/3300	0.73	0/4482
2	B	0.51	0/3426	0.76	2/4642 (0.0%)
3	C	0.54	0/1502	0.62	0/2027
3	E	0.54	0/1502	0.62	0/2027
4	D	0.61	0/1295	0.68	1/1751 (0.1%)
4	F	0.61	0/1295	0.68	1/1751 (0.1%)
All	All	0.54	0/12320	0.70	4/16680 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	235	MET	CG-SD-CE	6.09	109.95	100.20
2	B	217	LEU	N-CA-C	-5.36	96.53	111.00
4	D	31	LEU	CA-CB-CG	5.24	127.36	115.30
4	F	31	LEU	CA-CB-CG	5.22	127.31	115.30

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	3227	0	3143	551	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	3351	0	3228	553	0
3	C	1465	0	1432	46	0
3	E	1465	0	1432	27	0
4	D	1261	0	1263	33	0
4	F	1261	0	1263	51	0
5	A	1	0	0	0	0
6	A	1	0	0	0	0
7	A	32	0	12	4	0
8	B	28	0	12	1	0
9	B	62	0	51	5	0
All	All	12154	0	11836	1198	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 50.

The worst 5 of 1198 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:115:LYS:NZ	3:E:81:LYS:HZ3	1.04	1.49
3:C:115:LYS:HZ3	3:E:81:LYS:NZ	1.06	1.41
3:C:141:GLU:CD	4:F:29:LYS:HD2	1.03	1.38
3:C:141:GLU:OE1	4:F:29:LYS:CE	1.71	1.38
3:C:141:GLU:OE2	4:F:26:ALA:CA	1.85	1.24

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	408/451 (90%)	266 (65%)	83 (20%)	59 (14%)	0 6
2	B	424/445 (95%)	274 (65%)	94 (22%)	56 (13%)	0 7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	C	178/315 (56%)	167 (94%)	9 (5%)	2 (1%)	17 63
3	E	178/315 (56%)	167 (94%)	9 (5%)	2 (1%)	17 63
4	D	151/250 (60%)	144 (95%)	7 (5%)	0	100 100
4	F	151/250 (60%)	144 (95%)	7 (5%)	0	100 100
All	All	1490/2026 (74%)	1162 (78%)	209 (14%)	119 (8%)	2 19

5 of 119 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

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%)	4 26
2	B	367/381 (96%)	307 (84%)	60 (16%)	3 20
3	C	163/279 (58%)	155 (95%)	8 (5%)	31 67
3	E	163/279 (58%)	155 (95%)	8 (5%)	31 67
4	D	140/226 (62%)	126 (90%)	14 (10%)	9 38
4	F	140/226 (62%)	126 (90%)	14 (10%)	9 38
All	All	1320/1768 (75%)	1167 (88%)	153 (12%)	11 32

5 of 153 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	174	SER
2	B	275	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
4	F	24	THR
2	B	201	THR
2	B	224	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 45 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	331	GLN
2	B	406	HIS
4	F	129	HIS
2	B	337	ASN
2	B	436	GLN

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, 2 are monoatomic - leaving 3 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	A	920	6	26,34,34	1.35	1 (3%)	29,54,54	2.28	4 (13%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
8	GDP	B	800	-	24,30,30	2.65	8 (33%)	26,47,47	3.29	8 (30%)
9	TA1	B	820	-	68,68,68	1.93	20 (29%)	102,105,105	1.31	8 (7%)

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	A	920	6	-	0/18/38/38	0/3/3/3
8	GDP	B	800	-	-	0/12/32/32	0/3/3/3
9	TA1	B	820	-	-	0/41/127/127	0/5/7/7

The worst 5 of 29 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	B	820	TA1	C08-C07	-4.96	1.25	1.38
8	B	800	GDP	PB-O2B	-4.18	1.40	1.54
9	B	820	TA1	C04-C03	-2.32	1.44	1.49
9	B	820	TA1	C10-C02	2.02	1.62	1.57
9	B	820	TA1	C18-C20	2.04	1.62	1.56

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

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	B	800	GDP	C6-C5-C4	-9.93	109.51	120.86
7	A	920	GTP	C5-C6-N1	-7.77	113.36	123.52
8	B	800	GDP	N2-C2-N1	-5.69	107.81	117.20
8	B	800	GDP	N3-C2-N1	-5.35	120.27	127.56
9	B	820	TA1	C06-C05-C04	-4.84	114.58	120.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	920	GTP	4	0
8	B	800	GDP	1	0
9	B	820	TA1	5	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.