wwPDB EM Map/Model Validation Report (i)

Aug 29, 2016 – 11:57 AM EDT

PDB ID	:	5HNZ
EMDB ID:	:	EMD-8061
Title	:	Structural basis of backwards motion in kinesin-14: plus-end directed nKn669 in the nucleotide-free state
Authors	:	Shigematsu, H.; Yokoyama, T.; Kikkawa, M.; Shirouzu, M.; Nitta, R.
Deposited on	:	2016-01-19
Resolution	:	5.80 Å(reported)
Based on PDB ID	:	?

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

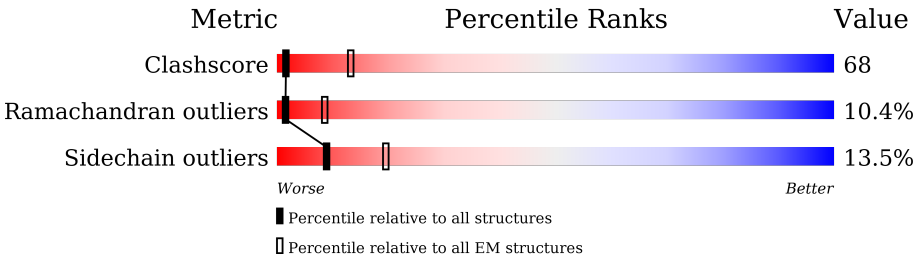
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027939

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 5.80 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	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% 18% • 9%
2	B	445	20% 56% 18% • •
3	K	371	60% 23% • 15%

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	502	-	-	X	-

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 9177 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-1B chain.

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

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	136	SER	LEU	conflict	UNP P81947
A	232	GLY	SER	conflict	UNP P81947
A	265	GLY	ILE	conflict	UNP P81947
A	340	THR	SER	conflict	UNP P81947
A	358	GLU	GLN	conflict	UNP P81947

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

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

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	57	ALA	THR	conflict	UNP Q6B856
B	172	VAL	MET	conflict	UNP Q6B856
B	298	ALA	SER	conflict	UNP Q6B856
B	318	VAL	ILE	conflict	UNP Q6B856

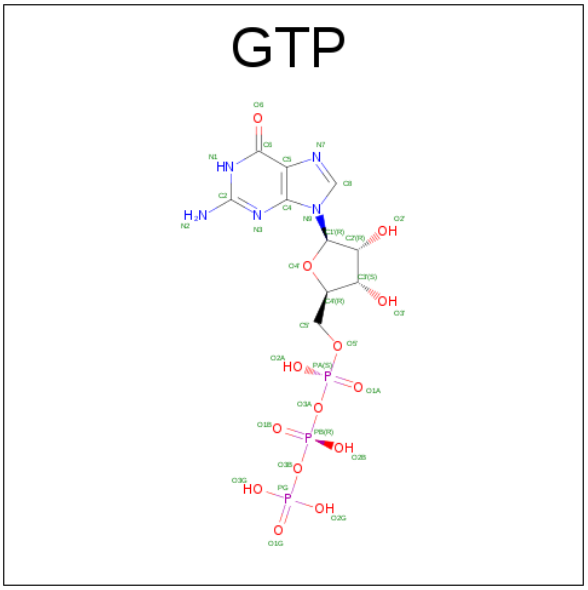
- Molecule 3 is a protein called Protein claret segregational,Protein claret segregational,Plus-end directed kinesin-1/kinesin-14,Protein claret segregational,Protein claret segregational.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	K	316	2476	1557	429	479	11	0	0

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

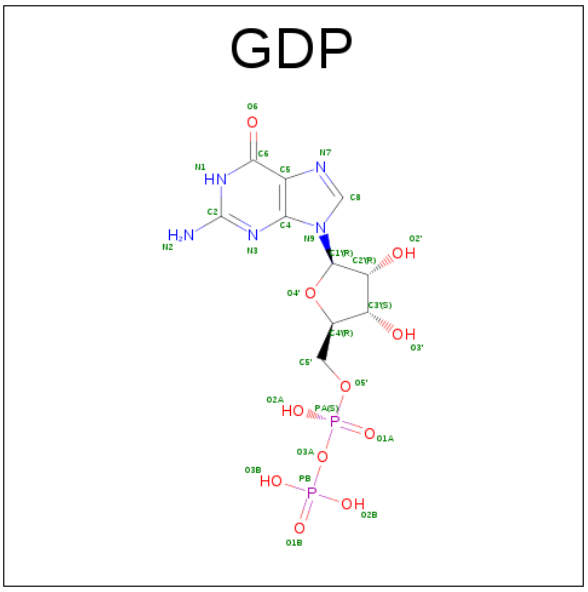
Mol	Chain	Residues	Atoms		AltConf
			Total	Mg	
4	A	1	1	1	0

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



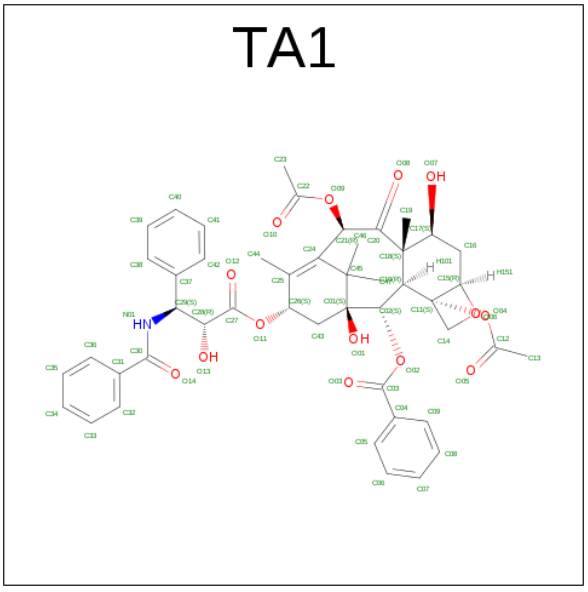
Mol	Chain	Residues	Atoms					AltConf
			Total	C	N	O	P	
5	A	1	32	10	5	14	3	0

- 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: $C_{47}H_{51}NO_{14}$).



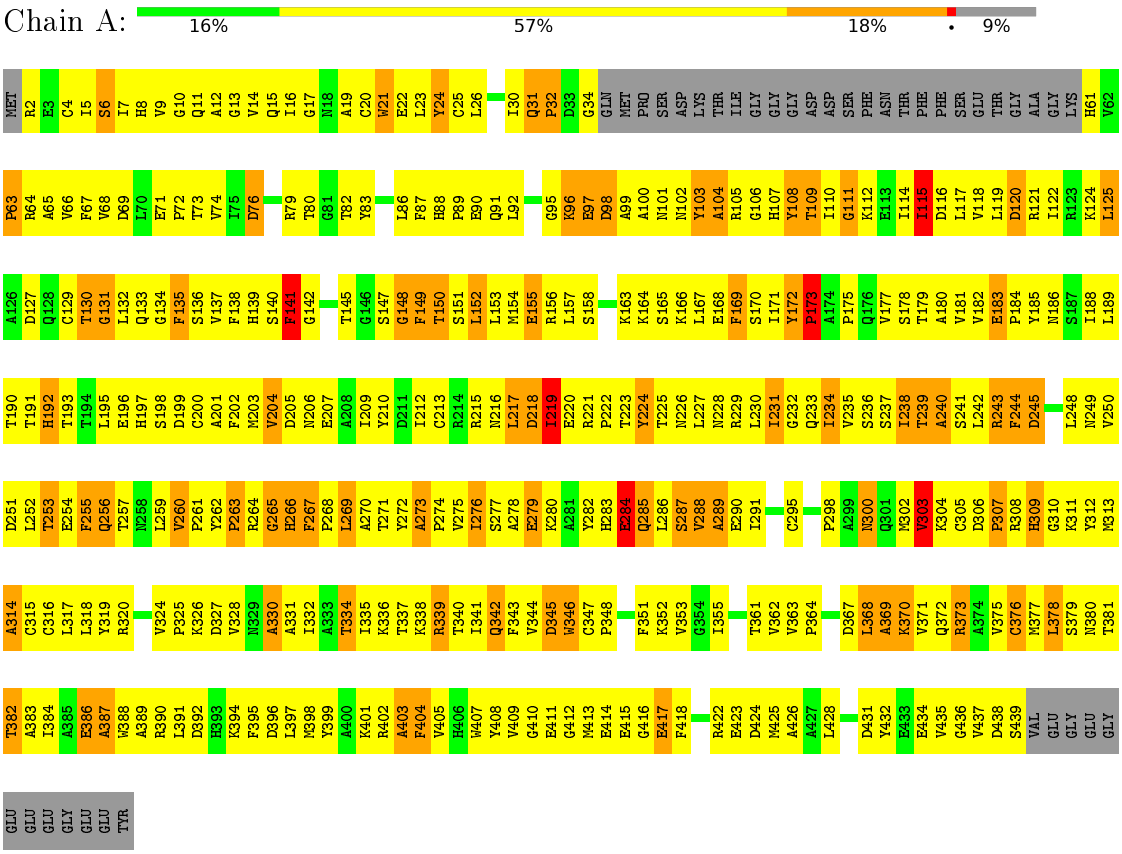
Mol	Chain	Residues	Atoms				AltConf
7	B	1	Total	C	N	O	0
			62	47	1	14	

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.

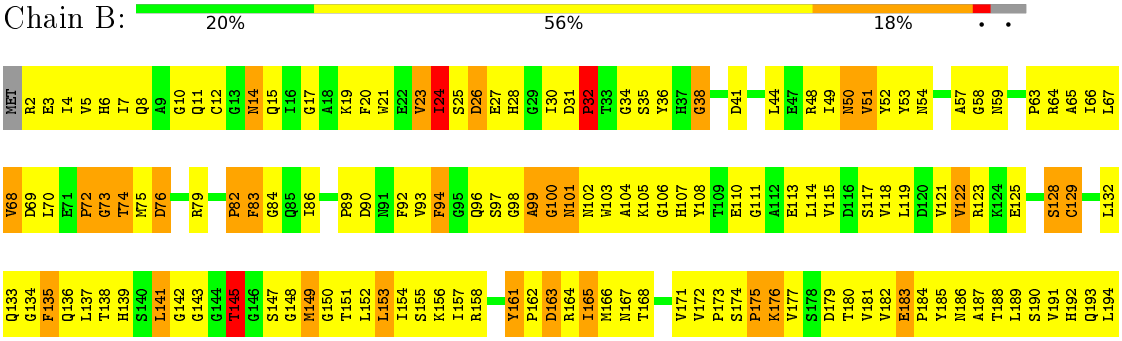
- Molecule 1: Tubulin alpha-1B chain

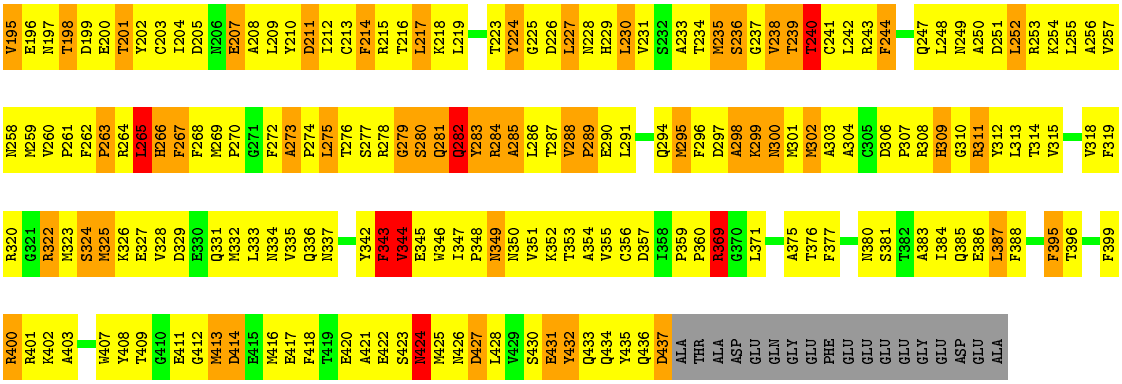
Chain A:



- Molecule 2: Tubulin beta-2B chain

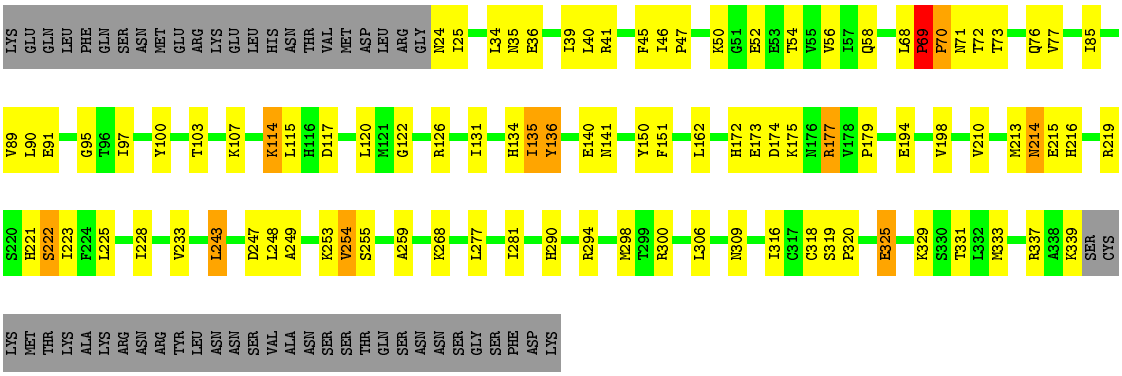
Chain B:





• Molecule 3: Protein claret segregational,Protein claret segregational,Plus-end directed kinesin-1 /kinesin-14,Protein claret segregational,Protein claret segregational

Chain K:



4 Experimental information (i)

Property	Value	Source
Reconstruction method	HELICAL	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of segments used	203826	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TECNAI ARCTICA	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/Å^2$)	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	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, TA1

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	0.51	0/3300	0.73	0/4482
2	B	0.51	0/3426	0.76	2/4642 (0.0%)
3	K	0.50	0/2517	0.76	2/3395 (0.1%)
All	All	0.51	0/9243	0.75	4/12519 (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	6.12	109.99	100.20
2	B	217	LEU	N-CA-C	-5.42	96.36	111.00
3	K	69	PRO	C-N-CD	5.09	139.09	128.40
3	K	319	SER	C-N-CD	5.04	138.98	128.40

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	614	0
2	B	3351	0	3229	625	0
3	K	2476	0	2481	104	0
4	A	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	32	0	12	12	0
6	B	28	0	12	2	0
7	B	62	0	51	10	0
All	All	9177	0	8928	1239	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 68.

The worst 5 of 1239 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:416:MET:CE	3:K:173:GLU:O	1.79	1.27
1:A:108:TYR:CE2	3:K:255:SER:HB2	1.68	1.26
1:A:101:ASN:ND2	2:B:254:LYS:HD2	1.58	1.19
2:B:423:SER:OG	3:K:177:ARG:NH2	1.76	1.17
1:A:11:GLN:N	5:A:502:GTP:O1B	1.78	1.15

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%)	265 (65%)	84 (21%)	59 (14%)	0 6
2	B	424/445 (95%)	274 (65%)	94 (22%)	56 (13%)	0 7
3	K	314/371 (85%)	292 (93%)	18 (6%)	4 (1%)	15 60
All	All	1146/1267 (90%)	831 (72%)	196 (17%)	119 (10%)	1 12

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 (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%)	4 26
2	B	367/381 (96%)	307 (84%)	60 (16%)	3 20
3	K	276/330 (84%)	251 (91%)	25 (9%)	12 44
All	All	990/1088 (91%)	856 (86%)	134 (14%)	9 27

5 of 134 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	129	CYS
2	B	214	PHE
3	K	177	ARG
2	B	141	LEU
2	B	165	ILE

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	102	ASN
2	B	197	ASN
3	K	134	HIS
2	B	136	GLN
1	A	197	HIS

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 4 ligands modelled in this entry, 1 is 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	$\#\|Z\| > 2$
5	GTP	A	502	4	26,34,34	1.34	1 (3%)	29,54,54	2.28	4 (13%)
6	GDP	B	901	-	24,30,30	2.69	8 (33%)	26,47,47	3.30	8 (30%)
7	TA1	B	902	-	68,68,68	1.94	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
5	GTP	A	502	4	-	0/18/38/38	0/3/3/3
6	GDP	B	901	-	-	0/12/32/32	0/3/3/3
7	TA1	B	902	-	-	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)
7	B	902	TA1	C08-C07	-4.99	1.25	1.38
6	B	901	GDP	PB-O2B	-4.23	1.40	1.54
7	B	902	TA1	C04-C03	-2.31	1.44	1.49
7	B	902	TA1	C10-C02	2.06	1.62	1.57
7	B	902	TA1	C18-C20	2.08	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})
6	B	901	GDP	C6-C5-C4	-9.90	109.54	120.86
5	A	502	GTP	C5-C6-N1	-7.70	113.46	123.52
6	B	901	GDP	N2-C2-N1	-5.74	107.73	117.20
6	B	901	GDP	N3-C2-N1	-5.38	120.23	127.56
7	B	902	TA1	C06-C05-C04	-4.79	114.63	120.35

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 24 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	502	GTP	12	0
6	B	901	GDP	2	0
7	B	902	TA1	10	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.