



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2018 – 12:07 am GMT

PDB ID : 5OOB  
Title : COMPLEX OF HUMAN NUCLEAR CAP-BINDING COMPLEX WITH  
M7GTP AND NELF-E C-TERMINAL PEPTIDE  
Authors : Cusack, S.; Schulze, W.M.  
Deposited on : 2017-08-07  
Resolution : 2.79 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk31020  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk31020

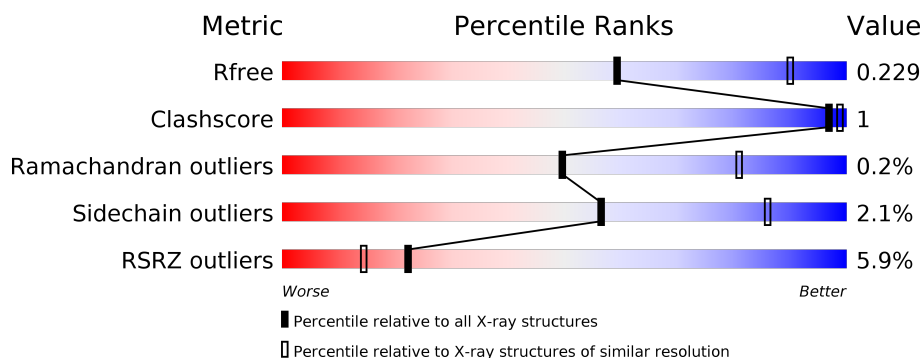
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.79 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)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	111664	2792 (2.80-2.80)
Clashscore	122126	3209 (2.80-2.80)
Ramachandran outliers	120053	3158 (2.80-2.80)
Sidechain outliers	120020	3160 (2.80-2.80)
RSRZ outliers	108989	2726 (2.80-2.80)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	772	<div> <div>5%</div> <div>91%</div> <div>5%</div> </div>
1	C	772	<div> <div>2%</div> <div>90%</div> <div>5%</div> </div>
1	F	772	<div> <div>10%</div> <div>88%</div> <div>5%</div> <div>7%</div> </div>
1	I	772	<div> <div>2%</div> <div>90%</div> <div>6%</div> </div>
2	B	158	<div> <div>3%</div> <div>91%</div> <div>6%</div> </div>
2	D	158	<div> <div>3%</div> <div>92%</div> <div>5%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	158	<div><div></div><div>13%</div><div>90%</div><div>6%</div></div>
2	J	158	<div><div></div><div>7%</div><div>89%</div><div>6%</div></div>
3	E	21	<div><div></div><div>19%</div><div>76%</div><div>10%</div><div>14%</div></div>
3	K	21	<div><div></div><div>29%</div><div>71%</div><div>10%</div><div>19%</div></div>
3	Z	21	<div><div></div><div>24%</div><div>62%</div><div>5%</div><div>33%</div></div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 29296 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, 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 Nuclear cap-binding protein subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	739	Total	C	N	O	S	0	0	0
			6045	3897	1018	1092	38			
1	C	731	Total	C	N	O	S	0	0	0
			5980	3853	1006	1083	38			
1	F	718	Total	C	N	O	S	0	0	0
			5879	3791	992	1058	38			
1	I	729	Total	C	N	O	S	0	0	0
			5964	3845	1005	1076	38			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	19	MET	-	initiating methionine	UNP Q09161
C	19	MET	-	initiating methionine	UNP Q09161
F	19	MET	-	initiating methionine	UNP Q09161
I	19	MET	-	initiating methionine	UNP Q09161

- Molecule 2 is a protein called Nuclear cap-binding protein subunit 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	152	Total	C	N	O	S	0	0	0
			1237	769	222	240	6			
2	D	151	Total	C	N	O	S	0	0	0
			1229	765	220	238	6			
2	G	149	Total	C	N	O	S	0	0	0
			1215	757	217	235	6			
2	J	148	Total	C	N	O	S	0	0	0
			1207	751	216	234	6			

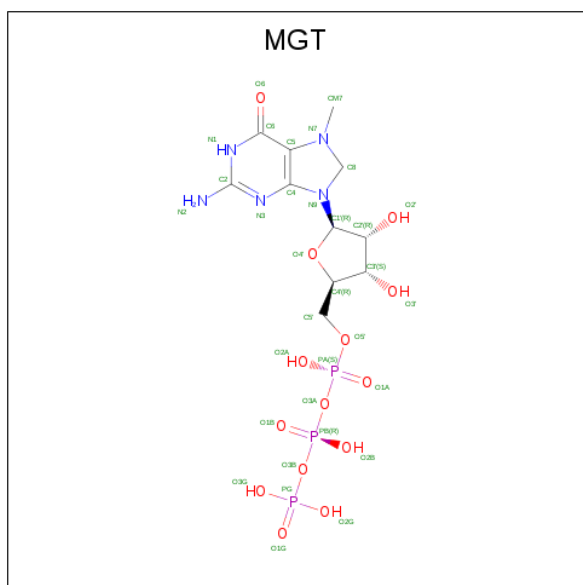
There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-1	GLY	-	expression tag	UNP P52298
B	0	ALA	-	expression tag	UNP P52298
D	-1	GLY	-	expression tag	UNP P52298
D	0	ALA	-	expression tag	UNP P52298
G	-1	GLY	-	expression tag	UNP P52298
G	0	ALA	-	expression tag	UNP P52298
J	-1	GLY	-	expression tag	UNP P52298
J	0	ALA	-	expression tag	UNP P52298

- Molecule 3 is a protein called Negative elongation factor E.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	E	18	Total	C	N	O	0	0	0
			150	93	24	33			
3	K	17	Total	C	N	O	0	0	0
			142	89	22	31			
3	Z	14	Total	C	N	O	0	0	0
			116	74	18	24			

- Molecule 4 is 7N-METHYL-8-HYDROGUANOSINE-5'-TRIPHOSPHATE (three-letter code: MGT) (formula:  $C_{11}H_{20}N_5O_{14}P_3$ ).



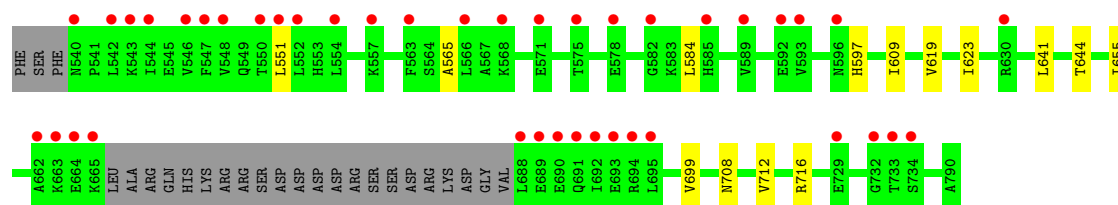
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	1	Total	C	N	O	0	0
			33	11	5	14		
4	D	1	Total	C	N	O	0	0
			33	11	5	14		

*Continued on next page...*

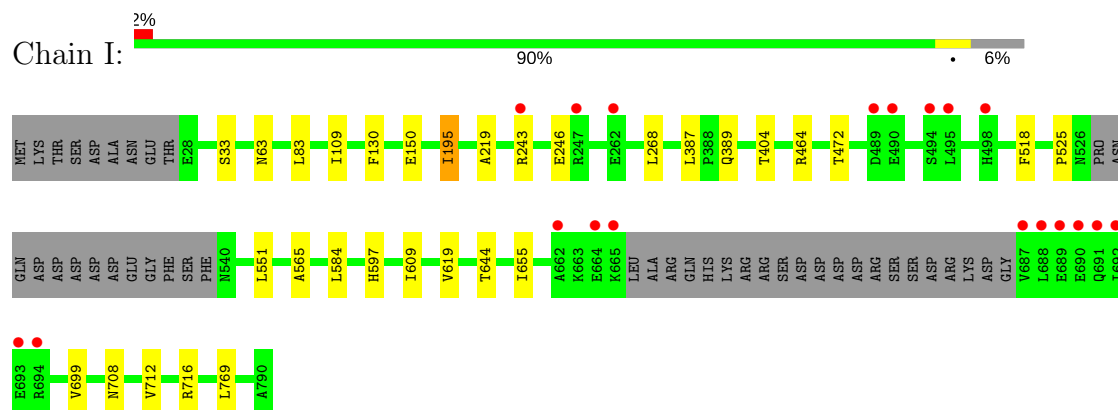
*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	G	1	Total	C	N	O	P	0	0
			33	11	5	14	3		
4	J	1	Total	C	N	O	P	0	0
			33	11	5	14	3		

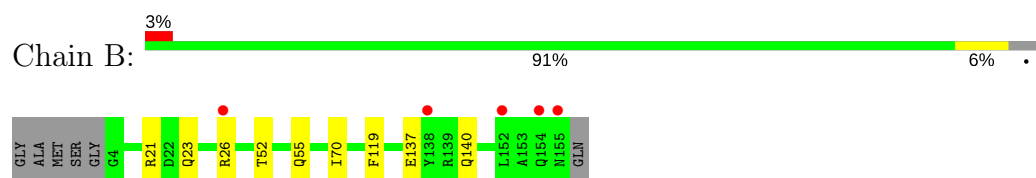




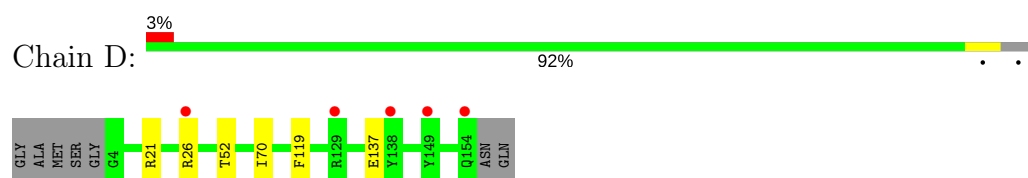
• Molecule 1: Nuclear cap-binding protein subunit 1



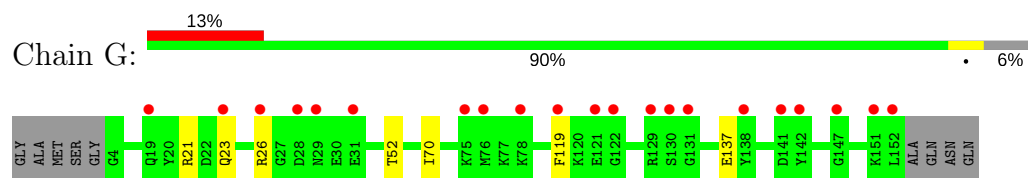
• Molecule 2: Nuclear cap-binding protein subunit 2



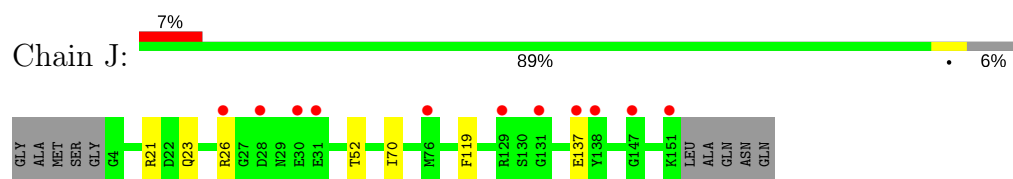
• Molecule 2: Nuclear cap-binding protein subunit 2



• Molecule 2: Nuclear cap-binding protein subunit 2

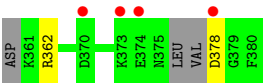
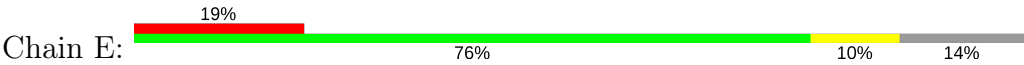


• Molecule 2: Nuclear cap-binding protein subunit 2

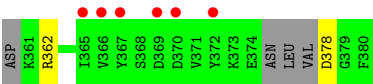


• Molecule 3: Negative elongation factor E

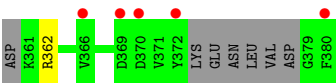




• Molecule 3: Negative elongation factor E



• Molecule 3: Negative elongation factor E



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	113.80Å 147.23Å 153.88Å 90.00° 91.48° 90.00°	Depositor
Resolution (Å)	153.83 – 2.79 49.79 – 2.79	Depositor EDS
% Data completeness (in resolution range)	98.3 (153.83-2.79) 98.4 (49.79-2.79)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.42 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.202 , 0.229 0.206 , 0.229	Depositor DCC
$R_{free}$ test set	6345 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	63.0	Xtriage
Anisotropy	0.518	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 35.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.002 for -h,l,k 0.012 for -h,-l,-k 0.023 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	29296	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.55% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 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: MGT

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	0.45	0/6198	0.62	0/8408
1	C	0.46	0/6132	0.63	0/8321
1	F	0.44	0/6026	0.63	0/8175
1	I	0.46	0/6115	0.63	0/8297
2	B	0.50	0/1257	0.74	0/1677
2	D	0.50	0/1249	0.72	0/1666
2	G	0.47	0/1235	0.72	0/1647
2	J	0.49	0/1227	0.73	0/1636
3	E	0.62	0/151	0.60	0/200
3	K	0.64	0/143	0.61	0/189
3	Z	0.63	0/117	0.60	0/155
All	All	0.46	0/29850	0.65	0/40371

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	6045	0	6028	16	0
1	C	5980	0	5952	13	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	F	5879	0	5876	16	0
1	I	5964	0	5954	12	0
2	B	1237	0	1193	2	0
2	D	1229	0	1187	1	0
2	G	1215	0	1174	1	0
2	J	1207	0	1163	1	0
3	E	150	0	129	0	0
3	K	142	0	123	0	0
3	Z	116	0	100	0	0
4	B	33	0	16	0	0
4	D	33	0	16	0	0
4	G	33	0	16	0	0
4	J	33	0	16	0	0
All	All	29296	0	28943	58	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 1.

The worst 5 of 58 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:C:609:ILE:HD11	1:C:619:VAL:HG21	1.63	0.81
1:F:609:ILE:HD11	1:F:619:VAL:HG21	1.61	0.80
1:I:609:ILE:HD11	1:I:619:VAL:HG21	1.62	0.79
1:A:609:ILE:HD11	1:A:619:VAL:HG21	1.63	0.77
1:A:659:LEU:HG	1:A:663:LYS:HE3	1.67	0.75

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 X-ray entries followed by that with respect to entries of similar resolution.

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	733/772 (95%)	715 (98%)	15 (2%)	3 (0%)	36	70
1	C	725/772 (94%)	709 (98%)	16 (2%)	0	100	100
1	F	710/772 (92%)	697 (98%)	13 (2%)	0	100	100
1	I	723/772 (94%)	707 (98%)	15 (2%)	1 (0%)	53	84
2	B	150/158 (95%)	140 (93%)	9 (6%)	1 (1%)	24	57
2	D	149/158 (94%)	139 (93%)	9 (6%)	1 (1%)	24	57
2	G	147/158 (93%)	138 (94%)	8 (5%)	1 (1%)	24	57
2	J	146/158 (92%)	135 (92%)	10 (7%)	1 (1%)	24	57
3	E	14/21 (67%)	12 (86%)	2 (14%)	0	100	100
3	K	13/21 (62%)	11 (85%)	2 (15%)	0	100	100
3	Z	10/21 (48%)	10 (100%)	0	0	100	100
All	All	3520/3783 (93%)	3413 (97%)	99 (3%)	8 (0%)	49	81

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	26	ARG
2	D	26	ARG
2	G	26	ARG
2	J	26	ARG
1	A	525	PRO

### 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 X-ray entries followed by that with respect to entries of similar resolution.

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	677/708 (96%)	664 (98%)	13 (2%)	60	87
1	C	670/708 (95%)	658 (98%)	12 (2%)	62	88
1	F	659/708 (93%)	647 (98%)	12 (2%)	62	88
1	I	669/708 (94%)	658 (98%)	11 (2%)	65	90
2	B	127/130 (98%)	122 (96%)	5 (4%)	35	69

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	D	126/130 (97%)	123 (98%)	3 (2%)	52 83
2	G	125/130 (96%)	121 (97%)	4 (3%)	42 76
2	J	124/130 (95%)	120 (97%)	4 (3%)	42 76
3	E	16/20 (80%)	14 (88%)	2 (12%)	5 15
3	K	15/20 (75%)	13 (87%)	2 (13%)	4 13
3	Z	12/20 (60%)	11 (92%)	1 (8%)	12 34
All	All	3220/3412 (94%)	3151 (98%)	69 (2%)	56 86

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	52	THR
1	F	246	GLU
2	J	23	GLN
2	D	137	GLU
1	F	33	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

4 ligands are modelled in this entry.

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
4	MGT	B	201	-	29,35,35	1.65	4 (13%)	34,56,56	2.22	7 (20%)
4	MGT	D	201	-	29,35,35	1.39	3 (10%)	34,56,56	2.34	7 (20%)
4	MGT	G	201	-	29,35,35	1.66	4 (13%)	34,56,56	2.19	7 (20%)
4	MGT	J	201	-	29,35,35	1.65	3 (10%)	34,56,56	2.26	7 (20%)

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
4	MGT	B	201	-	-	0/22/50/50	0/3/3/3
4	MGT	D	201	-	-	0/22/50/50	0/3/3/3
4	MGT	G	201	-	-	0/22/50/50	0/3/3/3
4	MGT	J	201	-	-	0/22/50/50	0/3/3/3

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	201	MGT	C4-N9	-4.07	1.33	1.38
4	D	201	MGT	C4-N9	-3.74	1.33	1.38
4	G	201	MGT	C4-N9	-3.47	1.33	1.38
4	J	201	MGT	C4-N9	-3.01	1.34	1.38
4	G	201	MGT	PG-O3B	2.02	1.63	1.60

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	201	MGT	C5-C6-N1	-5.41	114.88	123.34
4	B	201	MGT	C5-C6-N1	-5.26	115.12	123.34
4	J	201	MGT	C5-C4-N3	-5.16	117.85	126.47
4	J	201	MGT	C5-C6-N1	-5.04	115.46	123.34
4	B	201	MGT	C5-C4-N3	-4.83	118.41	126.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 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.



## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	739/772 (95%)	0.35	38 (5%) 28 19	41, 71, 117, 173	0
1	C	731/772 (94%)	0.17	18 (2%) 57 47	33, 59, 98, 141	0
1	F	718/772 (93%)	0.59	80 (11%) 5 3	43, 77, 144, 193	0
1	I	729/772 (94%)	0.20	19 (2%) 56 45	42, 66, 105, 144	0
2	B	152/158 (96%)	0.23	5 (3%) 46 36	43, 68, 110, 142	0
2	D	151/158 (95%)	0.19	5 (3%) 46 36	38, 68, 108, 127	0
2	G	149/158 (94%)	0.76	21 (14%) 2 1	59, 86, 118, 140	0
2	J	148/158 (93%)	0.55	11 (7%) 14 8	46, 79, 119, 144	0
3	E	18/21 (85%)	1.39	4 (22%) 0 0	84, 104, 133, 135	0
3	K	17/21 (80%)	1.76	6 (35%) 0 0	86, 114, 143, 143	0
3	Z	14/21 (66%)	1.58	5 (35%) 0 0	98, 107, 140, 157	0
All	All	3566/3783 (94%)	0.36	212 (5%) 22 14	33, 70, 121, 193	0

The worst 5 of 212 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	688	LEU	7.8
1	A	666	LEU	7.7
1	F	521	LEU	7.6
1	F	518	PHE	7.3
1	F	544	ILE	6.3

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

### 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	MGT	J	201	33/33	0.95	0.14	57,74,91,105	0
4	MGT	G	201	33/33	0.96	0.14	68,76,88,102	0
4	MGT	D	201	33/33	0.97	0.14	52,60,73,82	0
4	MGT	B	201	33/33	0.97	0.15	55,64,74,79	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.