



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 12:39 am GMT

PDB ID : 3M1I  
Title : Crystal structure of yeast CRM1 (Xpo1p) in complex with yeast RanBP1 (Yrb1p) and yeast RanGTP (Gsp1pGTP)  
Authors : Koyama, M.; Matsuura, Y.  
Deposited on : 2010-03-05  
Resolution : 2.00 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

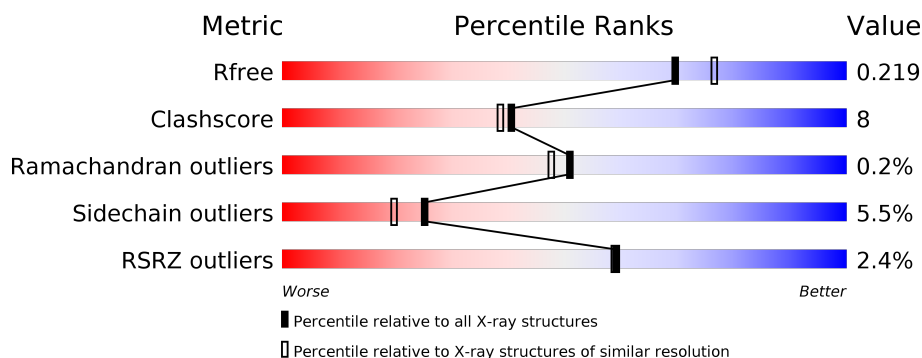
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	6609 (2.00-2.00)
Clashscore	112137	7775 (2.00-2.00)
Ramachandran outliers	110173	7679 (2.00-2.00)
Sidechain outliers	110143	7678 (2.00-2.00)
RSRZ outliers	101464	6696 (2.00-2.00)

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	219	<div> <div>3%</div> <div> <div></div> <div>78%</div> <div>10%</div> <div>•</div> <div>9%</div> </div> </div>
2	B	191	<div> <div>2%</div> <div> <div></div> <div>62%</div> <div>•</div> <div>•</div> <div>31%</div> </div> </div>
3	C	1049	<div> <div>2%</div> <div> <div></div> <div>78%</div> <div>16%</div> <div>•</div> <div>•</div> </div> </div>

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 11634 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 GTP-binding nuclear protein GSP1/CNR1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	200	Total	C	N	O	S	0	0	0
			1596	1033	268	290	5			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	71	LEU	GLN	ENGINEERED MUTATION	UNP P32835

- Molecule 2 is a protein called Ran-specific GTPase-activating protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	131	Total	C	N	O	S	0	0	0
			1055	668	184	198	5			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	MET	DELETION	UNP P41920
B	?	-	SER	DELETION	UNP P41920
B	?	-	SER	DELETION	UNP P41920
B	?	-	GLU	DELETION	UNP P41920
B	?	-	ASP	DELETION	UNP P41920
B	?	-	LYS	DELETION	UNP P41920
B	?	-	LYS	DELETION	UNP P41920
B	?	-	PRO	DELETION	UNP P41920
B	?	-	VAL	DELETION	UNP P41920
B	?	-	VAL	DELETION	UNP P41920

- Molecule 3 is a protein called Exportin-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	1023	Total	C	N	O	S	0	0	0
			8181	5261	1343	1538	39			

There are 39 discrepancies between the modelled and reference sequences:

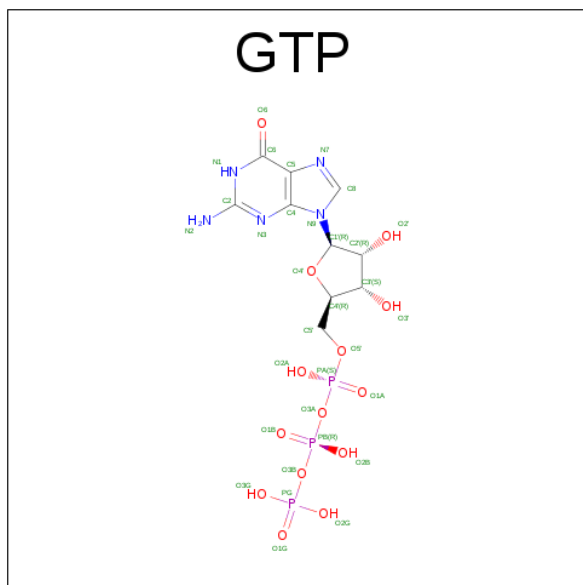
Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	GLY	-	EXPRESSION TAG	UNP P30822
C	0	ALA	-	EXPRESSION TAG	UNP P30822
C	?	-	VAL	DELETION	UNP P30822
C	?	-	GLN	DELETION	UNP P30822
C	?	-	ARG	DELETION	UNP P30822
C	?	-	LEU	DELETION	UNP P30822
C	?	-	PRO	DELETION	UNP P30822
C	?	-	ALA	DELETION	UNP P30822
C	?	-	THR	DELETION	UNP P30822
C	?	-	GLU	DELETION	UNP P30822
C	?	-	MET	DELETION	UNP P30822
C	?	-	SER	DELETION	UNP P30822
C	?	-	PRO	DELETION	UNP P30822
C	?	-	LEU	DELETION	UNP P30822
C	?	-	ILE	DELETION	UNP P30822
C	?	-	GLN	DELETION	UNP P30822
C	?	-	LEU	DELETION	UNP P30822
C	?	-	SER	DELETION	UNP P30822
C	?	-	VAL	DELETION	UNP P30822
C	?	-	GLY	DELETION	UNP P30822
C	?	-	SER	DELETION	UNP P30822
C	?	-	GLN	DELETION	UNP P30822
C	?	-	ALA	DELETION	UNP P30822
C	?	-	ILE	DELETION	UNP P30822
C	?	-	SER	DELETION	UNP P30822
C	?	-	THR	DELETION	UNP P30822
C	?	-	GLY	DELETION	UNP P30822
C	?	-	SER	DELETION	UNP P30822
C	?	-	GLY	DELETION	UNP P30822
C	?	-	ALA	DELETION	UNP P30822
C	?	-	LEU	DELETION	UNP P30822
C	?	-	ASN	DELETION	UNP P30822
C	?	-	PRO	DELETION	UNP P30822
C	?	-	GLU	DELETION	UNP P30822
C	?	-	TYR	DELETION	UNP P30822
C	?	-	MET	DELETION	UNP P30822
C	?	-	LYS	DELETION	UNP P30822

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	?	-	ARG	DELETION	UNP P30822
C	?	-	PHE	DELETION	UNP P30822

- Molecule 4 is GUANOSINE-5'-TRIPHOSPHATE (three-letter code: GTP) (formula:  $C_{10}H_{16}N_5O_{14}P_3$ ).



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

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

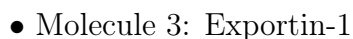
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total	Mg		
			1	1	0	0

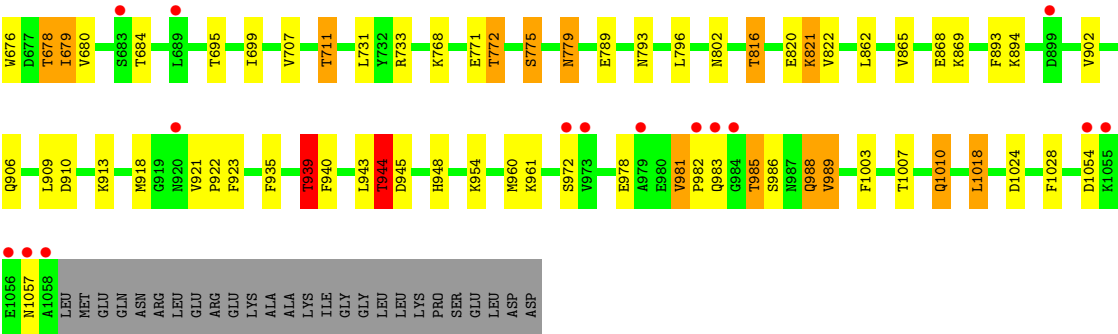
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	151	Total	O		
			151	151	0	0
6	B	45	Total	O		
			45	45	0	0
6	C	573	Total	O		
			573	573	0	0



- Molecule 1: GTP-binding nuclear protein GSP1/CNR1





E1056	0.00
M1057	0.00
A1058	0.00
LEU	0.00
MET	0.00
GLU	0.00
GLN	0.00
ASN	0.00
ARG	0.00
LEU	0.00
GLU	0.00
ARG	0.00
GLU	0.00
LYS	0.00
ALA	0.00
ALA	0.00
LYS	0.00
ILE	0.00
GLY	0.00
GLY	0.00
LEU	0.00
LEU	0.00
LYS	0.00
PRO	0.00
SER	0.00
GLU	0.00
LEU	0.00
ASP	0.00
ASP	0.00

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	106.21 Å   106.21 Å   303.60 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	43.51 – 2.00 43.51 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (43.51-2.00) 99.9 (43.51-2.00)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	0.08	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.20 (at 2.00 Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.175   ,   0.220 0.174   ,   0.219	Depositor DCC
$R_{free}$ test set	5918 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.7	Xtriage
Anisotropy	0.068	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 45.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	11634	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	41.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.15% 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: GTP, 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	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	1.27	3/1636 (0.2%)	1.13	6/2219 (0.3%)
2	B	1.02	1/1072 (0.1%)	1.06	6/1433 (0.4%)
3	C	1.17	13/8339 (0.2%)	1.02	28/11317 (0.2%)
All	All	1.17	17/11047 (0.2%)	1.04	40/14969 (0.3%)

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	1	0
2	B	0	1
3	C	1	1
All	All	2	2

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	440	GLU	CB-CG	-6.97	1.38	1.52
3	C	1010	GLN	CB-CG	-6.39	1.35	1.52
2	B	90	ARG	CB-CG	-6.39	1.35	1.52
3	C	775	SER	CB-OG	-6.09	1.34	1.42
3	C	165	GLU	CB-CG	6.08	1.63	1.52

The worst 5 of 40 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	31	ARG	NE-CZ-NH1	16.35	128.47	120.30
3	C	219	ARG	NE-CZ-NH1	13.44	127.02	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	177	ARG	NE-CZ-NH2	-10.92	114.84	120.30
1	A	91	MET	CG-SD-CE	10.70	117.31	100.20
3	C	91	ARG	NE-CZ-NH2	-10.57	115.02	120.30

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	95	THR	CB
3	C	275	THR	CB

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	76	VAL	Peptide
3	C	204	GLY	Peptide

## 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	1596	0	1573	17	0
2	B	1055	0	1050	10	0
3	C	8181	0	8201	153	0
4	A	32	0	12	0	0
5	A	1	0	0	0	0
6	A	151	0	0	3	0
6	B	45	0	0	0	0
6	C	573	0	0	29	0
All	All	11634	0	10836	177	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 8.

The worst 5 of 177 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:43:ASP:HB2	6:C:1551:HOH:O	1.53	1.08
2:B:170:GLU:HG3	2:B:172:PHE:CE2	1.89	1.07
3:C:76:THR:HG22	6:C:1456:HOH:O	1.56	1.05
1:A:217:ALA:HB1	2:B:154:ARG:HH22	1.34	0.92
3:C:101:MET:HE2	6:C:1565:HOH:O	1.68	0.92

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	196/219 (90%)	191 (97%)	4 (2%)	1 (0%)	32	26
2	B	129/191 (68%)	121 (94%)	8 (6%)	0	100	100
3	C	1021/1049 (97%)	992 (97%)	27 (3%)	2 (0%)	51	48
All	All	1346/1459 (92%)	1304 (97%)	39 (3%)	3 (0%)	51	48

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	216	ASP
3	C	978	GLU
3	C	983	GLN

### 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 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	169/188 (90%)	161 (95%)	8 (5%)	30	26
2	B	108/164 (66%)	103 (95%)	5 (5%)	31	27
3	C	914/955 (96%)	862 (94%)	52 (6%)	24	18
All	All	1191/1307 (91%)	1126 (94%)	65 (6%)	25	20

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

Mol	Chain	Res	Type
3	C	262	LEU
3	C	425	SER
3	C	961	LYS
3	C	266	GLN
3	C	291	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	446	ASN
3	C	477	HIS
3	C	988	GLN
3	C	173	GLN
3	C	349	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 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 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	$\# Z  > 2$
4	GTP	A	1177	5	27,34,34	1.31	2 (7%)	27,54,54	2.16	9 (33%)

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	GTP	A	1177	5	-	0/18/38/38	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	1177	GTP	PG-O3B	2.69	1.64	1.60
4	A	1177	GTP	C2-N1	4.12	1.42	1.35

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	1177	GTP	N3-C2-N1	-5.66	119.20	127.46
4	A	1177	GTP	C5-C6-N1	-3.11	119.06	123.48
4	A	1177	GTP	O3'-C3'-C2'	-2.72	103.11	111.83
4	A	1177	GTP	C6-C5-C4	-2.72	118.14	120.84
4	A	1177	GTP	O3B-PG-O1G	-2.19	97.96	111.44

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 [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

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	200/219 (91%)	0.05	6 (3%)	51	51	20, 33, 68, 95	0
2	B	131/191 (68%)	-0.07	4 (3%)	49	49	31, 47, 80, 110	0
3	C	1023/1049 (97%)	-0.08	22 (2%)	62	61	20, 39, 64, 102	0
All	All	1354/1459 (92%)	-0.06	32 (2%)	59	59	20, 39, 67, 110	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	76	VAL	5.4
3	C	1057	ASN	5.0
3	C	1055	LYS	4.8
3	C	982	PRO	4.3
2	B	75	ASP	4.3

### 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	GTP	A	1177	32/32	0.98	0.11	-0.22	21,26,31,34	0
5	MG	A	1178	1/1	1.00	0.04	-3.73	18,18,18,18	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.