



# Full wwPDB X-ray Structure Validation Report ⓘ

Mar 9, 2018 – 01:50 am GMT

PDB ID : 5NQT  
Title : Tubulin Darpin room-temperature structure determined by serial millisecond crystallography  
Authors : Weinert, T.; Olieric, N.; James, D.; Gashi, D.; Nogly, P.; Jaeger, K.; Steinmetz, M.O.; Standfuss, J.  
Deposited on : 2017-04-21  
Resolution : 2.15 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : trunk30967  
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) : trunk30967

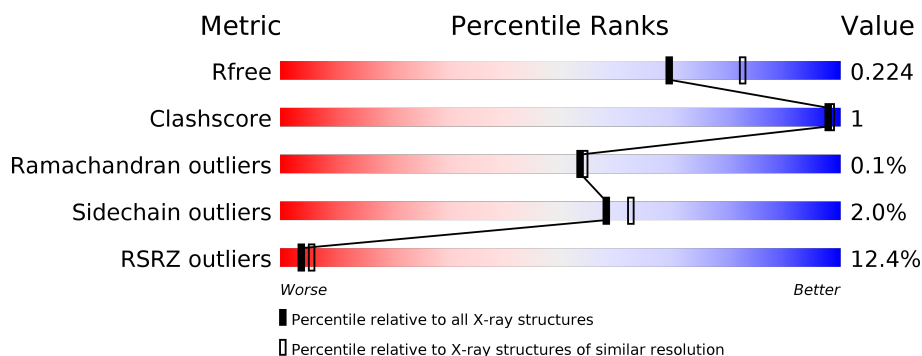
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.15 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	1287 (2.16-2.16)
Clashscore	122126	1390 (2.16-2.16)
Ramachandran outliers	120053	1368 (2.16-2.16)
Sidechain outliers	120020	1367 (2.16-2.16)
RSRZ outliers	108989	1262 (2.16-2.16)

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	451	<div> <div>10%</div> <div>93%</div> <div>5%</div> <div>• •</div> </div>
2	B	445	<div> <div>17%</div> <div>91%</div> <div>5%</div> <div>•</div> </div>
3	F	169	<div> <div>2%</div> <div>89%</div> <div>8%</div> <div>•</div> </div>

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 15755 atoms, of which 7733 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 Tubulin alpha-1B chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	431	Total	C	H	N	O	S	4	9	0
			6735	2162	3329	579	640	25			

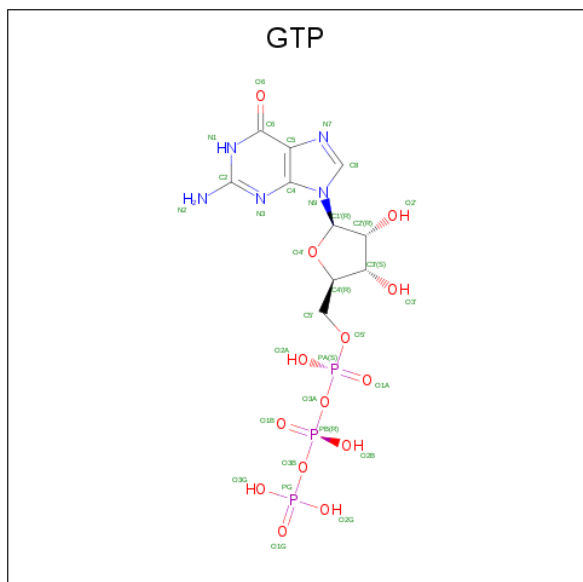
- Molecule 2 is a protein called Tubulin beta-2B chain.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	431	Total	C	H	N	O	S	0	4	0
			6600	2117	3229	573	654	27			

- Molecule 3 is a protein called DESIGNED ANKYRIN REPEAT PROTEIN (DARPIN) D1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	F	155	Total	C	H	N	O	S	0	1	0
			2305	728	1151	195	228	3			

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

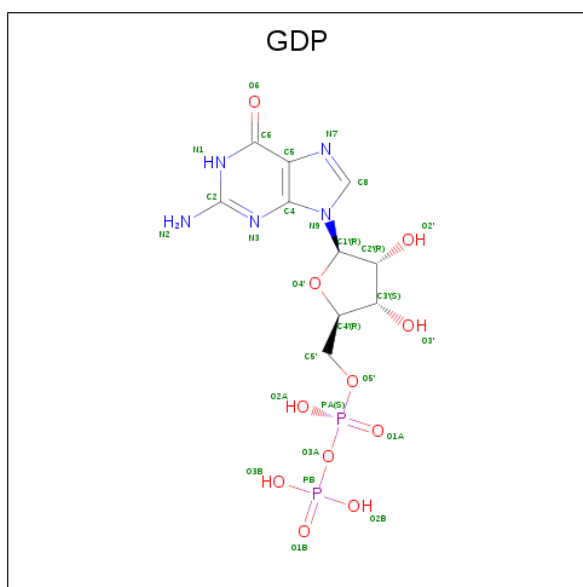


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	H	N	O	P	
			44	10	12	5	14	3	
								0	0

- 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 GUANOSINE-5'-DIPHOSPHATE (three-letter code: GDP) (formula: C<sub>10</sub>H<sub>15</sub>N<sub>5</sub>O<sub>11</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	B	1	Total	C	H	N	O	P	
			40	10	12	5	11	2	
								0	0

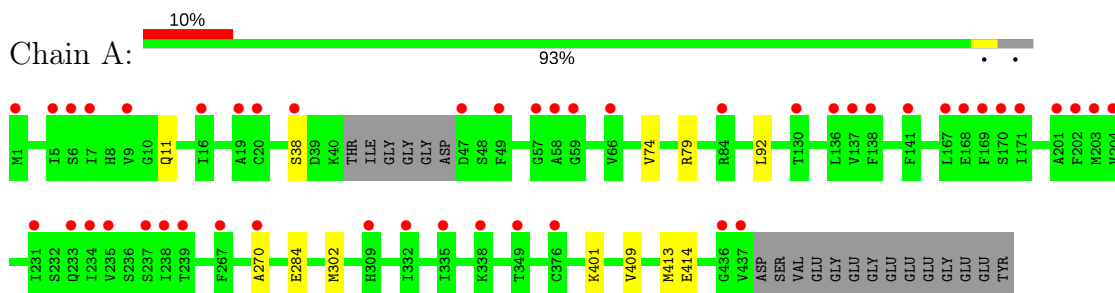
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	17	Total	O		
			17	17	0	0
7	B	8	Total	O		
			8	8	0	0
7	F	5	Total	O		
			5	5	0	0

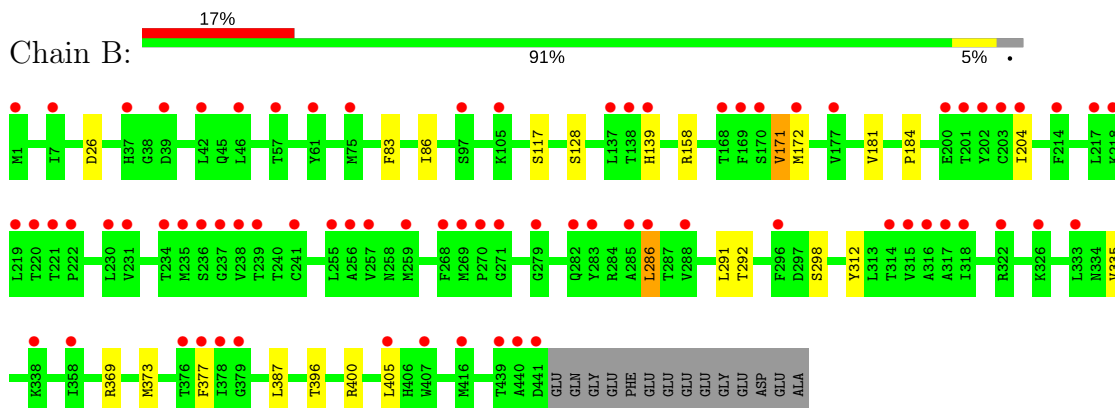
### 3 Residue-property plots

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 and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). 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.

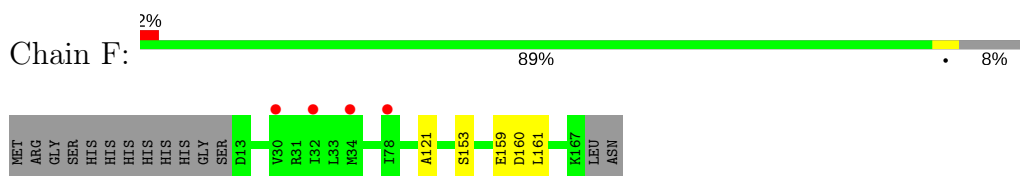
#### • Molecule 1: Tubulin alpha-1B chain



#### • Molecule 2: Tubulin beta-2B chain



#### • Molecule 3: DESIGNED ANKYRIN REPEAT PROTEIN (DARPIN) D1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.04Å 91.43Å 83.61Å 90.00° 96.82° 90.00°	Depositor
Resolution (Å)	33.60 – 2.15 32.84 – 2.15	Depositor EDS
% Data completeness (in resolution range)	99.9 (33.60-2.15) 99.9 (32.84-2.15)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.25 (at 2.16Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, $R_{free}$	0.179 , 0.226 0.185 , 0.224	Depositor DCC
$R_{free}$ test set	2983 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	55.5	Xtriage
Anisotropy	0.188	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 62.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	15755	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 4.74% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GDP, 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	0.54	0/3513	0.67	0/4767
2	B	0.48	0/3458	0.63	0/4687
3	F	0.51	0/1173	0.59	0/1595
All	All	0.51	0/8144	0.64	0/11049

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

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	3406	3329	3314	4	0
2	B	3371	3229	3220	8	0
3	F	1154	1151	1151	1	0
4	A	32	12	12	0	0
5	A	1	0	0	0	0
6	B	28	12	12	0	0
7	A	17	0	0	0	0
7	B	8	0	0	0	0
7	F	5	0	0	0	0
All	All	8022	7733	7709	13	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.

All (13) 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:292:THR:HG22	2:B:335:VAL:HG21	1.88	0.56
2:B:172:MET:HG3	2:B:387:LEU:HD21	1.87	0.55
2:B:83:PHE:O	2:B:86:ILE:HG22	2.09	0.53
2:B:171:VAL:HA	2:B:204:ILE:O	2.12	0.49
2:B:396:THR:O	2:B:400:ARG:HG3	2.11	0.49
1:A:409:VAL:HA	1:A:413:MET:O	2.17	0.45
2:B:312:TYR:CE1	2:B:377:PHE:HZ	2.35	0.44
1:A:11:GLN:HG3	1:A:74:VAL:HG21	2.00	0.44
2:B:286:LEU:HD12	2:B:291:LEU:HG	2.00	0.43
2:B:181:VAL:O	2:B:184:PRO:HD2	2.19	0.43
3:F:121:ALA:HB1	3:F:161:LEU:HD21	2.01	0.43
1:A:79:ARG:HG2	1:A:92:LEU:HD13	2.01	0.42
1:A:270:ALA:HB3	1:A:302[A]:MET:HE2	2.02	0.41

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	436/451 (97%)	421 (97%)	14 (3%)	1 (0%)	49	48
2	B	433/445 (97%)	424 (98%)	9 (2%)	0	100	100
3	F	154/169 (91%)	153 (99%)	1 (1%)	0	100	100
All	All	1023/1065 (96%)	998 (98%)	24 (2%)	1 (0%)	53	54

All (1) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	A	38	SER

### 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	368/379 (97%)	365 (99%)	3 (1%)	83	87
2	B	367/383 (96%)	355 (97%)	12 (3%)	41	40
3	F	120/132 (91%)	117 (98%)	3 (2%)	50	53
All	All	855/894 (96%)	837 (98%)	18 (2%)	58	60

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	284	GLU
1	A	401	LYS
1	A	414	GLU
2	B	26	ASP
2	B	117	SER
2	B	128	SER
2	B	139	HIS
2	B	158[A]	ARG
2	B	158[B]	ARG
2	B	171	VAL
2	B	286	LEU
2	B	298	SER
2	B	369	ARG
2	B	373	MET
2	B	405	LEU
3	F	153	SER
3	F	159	GLU
3	F	160	ASP

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](#)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 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	501	5	27,34,34	0.80	0	26,54,54	0.82	0
6	GDP	B	501	-	25,30,30	0.81	0	24,47,47	1.13	3 (12%)

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	501	5	-	0/18/38/38	0/3/3/3
6	GDP	B	501	-	-	0/12/32/32	0/3/3/3

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	B	501	GDP	PA-O3A-PB	-2.87	122.97	132.63
6	B	501	GDP	C2-N3-C4	-2.16	112.63	115.16
6	B	501	GDP	O3B-PB-O2B	2.67	118.17	107.59

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

### 6.1 Protein, DNA and RNA chains

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	431/451 (95%)	0.54	47 (10%) <b>5</b> <b>8</b>	47, 65, 98, 153	3 (0%)
2	B	431/445 (96%)	0.82	75 (17%) <b>1</b> <b>1</b>	53, 77, 118, 151	0
3	F	155/169 (91%)	0.14	4 (2%) 56 63	52, 67, 94, 151	0
All	All	1017/1065 (95%)	0.60	126 (12%) <b>4</b> <b>5</b>	47, 70, 109, 153	3 (0%)

All (126) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	437	VAL	12.5
2	B	57	THR	5.6
2	B	42	LEU	5.5
1	A	169	PHE	5.4
2	B	1	MET	5.4
1	A	238	ILE	5.3
2	B	202	TYR	5.2
1	A	234	ILE	5.1
2	B	440	ALA	4.9
1	A	349	THR	4.7
1	A	235	VAL	4.6
2	B	285	ALA	4.6
2	B	268	PHE	4.6
1	A	1	MET	4.6
2	B	169	PHE	4.3
1	A	170	SER	4.3
2	B	172	MET	4.2
2	B	238	VAL	4.2
1	A	202	PHE	4.2
2	B	283	TYR	4.2
2	B	37	HIS	4.0
3	F	30	VAL	4.0
2	B	138	THR	4.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	441	ASP	3.9
2	B	378	ILE	3.9
1	A	138	PHE	3.8
2	B	168	THR	3.7
2	B	317	ALA	3.7
2	B	377	PHE	3.7
1	A	49	PHE	3.7
2	B	255	LEU	3.7
2	B	256	ALA	3.7
1	A	231	ILE	3.6
1	A	57	GLY	3.6
2	B	333	LEU	3.6
1	A	20	CYS	3.6
2	B	170	SER	3.6
1	A	309	HIS	3.5
2	B	137	LEU	3.5
1	A	47	ASP	3.5
2	B	270	PRO	3.5
2	B	379	GLY	3.5
1	A	204	VAL	3.4
1	A	59	GLY	3.3
2	B	259	MET	3.3
1	A	237	SER	3.2
2	B	416	MET	3.2
1	A	137	VAL	3.2
2	B	61	TYR	3.2
1	A	338	LYS	3.2
1	A	136	LEU	3.1
1	A	171	ILE	3.1
2	B	439	THR	3.1
2	B	219	LEU	3.1
1	A	167	LEU	3.1
2	B	75	MET	3.1
1	A	38	SER	3.0
1	A	376	CYS	3.0
2	B	358	ILE	3.0
2	B	271	GLY	3.0
1	A	270	ALA	3.0
2	B	204	ILE	3.0
2	B	220	THR	2.9
2	B	286	LEU	2.8
2	B	235	MET	2.8

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	46	LEU	2.8
1	A	16	ILE	2.8
1	A	203	MET	2.8
1	A	168	GLU	2.8
2	B	139	HIS	2.8
2	B	269	MET	2.8
1	A	66	VAL	2.7
2	B	201	THR	2.7
2	B	322	ARG	2.7
2	B	203	CYS	2.7
2	B	257	VAL	2.7
1	A	201	ALA	2.7
2	B	316	ALA	2.6
1	A	239	THR	2.6
2	B	405	LEU	2.6
1	A	58	ALA	2.6
2	B	241	CYS	2.6
2	B	218	LYS	2.6
2	B	231	VAL	2.5
1	A	7	ILE	2.5
1	A	141	PHE	2.5
1	A	332	ILE	2.4
1	A	130	THR	2.4
2	B	237	GLY	2.4
2	B	239	THR	2.4
2	B	318	ILE	2.4
3	F	32	ILE	2.4
2	B	279	GLY	2.4
2	B	338	LYS	2.4
2	B	326	LYS	2.3
2	B	221	THR	2.3
1	A	335	ILE	2.3
2	B	200	GLU	2.3
2	B	222	PRO	2.3
2	B	236	SER	2.3
2	B	234	THR	2.3
1	A	9	VAL	2.3
3	F	34	MET	2.3
1	A	6	SER	2.3
2	B	39	ASP	2.2
2	B	214	PHE	2.2
2	B	177	VAL	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	105	LYS	2.2
2	B	314	THR	2.2
3	F	78	ILE	2.2
2	B	296	PHE	2.2
2	B	282	GLN	2.1
2	B	315	VAL	2.1
1	A	19	ALA	2.1
1	A	436	GLY	2.1
2	B	288	VAL	2.1
1	A	267	PHE	2.1
2	B	376	THR	2.1
1	A	84	ARG	2.1
2	B	7	ILE	2.0
2	B	217	LEU	2.0
2	B	230	LEU	2.0
2	B	97	SER	2.0
1	A	233	GLN	2.0
2	B	407	TRP	2.0
1	A	5	ILE	2.0

## 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
6	GDP	B	501	28/28	0.96	0.13	62,67,74,74	0
5	MG	A	502	1/1	0.97	0.28	57,57,57,57	0
4	GTP	A	501	32/32	0.98	0.17	44,51,55,57	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.