



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 19, 2023 – 05:36 PM JST

PDB ID : 6LTS  
Title : Crystal structure of *Thermus thermophilus* transcription initiation complex comprising a truncated sigma finger  
Authors : Zhang, Y.; Ebright, R.H.  
Deposited on : 2020-01-23  
Resolution : 3.45 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

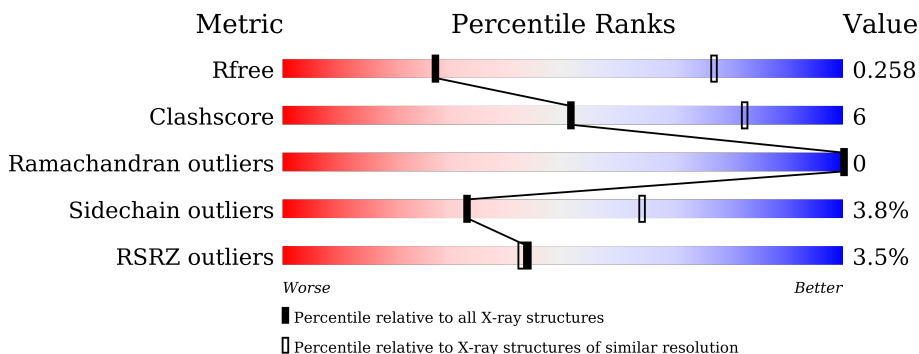
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.45 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))
$R_{free}$	130704	1291 (3.52-3.40)
Clashscore	141614	1372 (3.52-3.40)
Ramachandran outliers	138981	1337 (3.52-3.40)
Sidechain outliers	138945	1338 (3.52-3.40)
RSRZ outliers	127900	1205 (3.52-3.40)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. 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	315	58% 14% 28%
1	B	315	56% 13% 30%
2	C	1119	2% 82% 16% ..
3	D	1524	5% 81% 16% ..
4	E	99	2% 82% 12% 5%
5	F	443	3% 63% 12% 24%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
6	H	27	
7	G	19	

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
8	MG	B	2001	-	-	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 28227 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 DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	226	Total	C	N	O	S	0	0	0
			1782	1138	310	332	2			
1	B	222	Total	C	N	O	S	0	0	0
			1750	1118	304	326	2			

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	1111	Total	C	N	O	S	0	0	0
			8770	5548	1564	1634	24			

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	1486	Total	C	N	O	S	0	0	0
			11738	7441	2067	2195	35			

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	E	94	Total	C	N	O	S	0	0	0
			758	483	132	139	4			

- Molecule 5 is a protein called RNA polymerase sigma factor SigA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	F	336	Total	C	N	O	S	0	0	0
			2725	1717	498	506	4			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-19	MET	-	initiating methionine	UNP Q5SKW1
F	-18	GLY	-	expression tag	UNP Q5SKW1
F	-17	SER	-	expression tag	UNP Q5SKW1
F	-16	SER	-	expression tag	UNP Q5SKW1
F	-15	HIS	-	expression tag	UNP Q5SKW1
F	-14	HIS	-	expression tag	UNP Q5SKW1
F	-13	HIS	-	expression tag	UNP Q5SKW1
F	-12	HIS	-	expression tag	UNP Q5SKW1
F	-11	HIS	-	expression tag	UNP Q5SKW1
F	-10	HIS	-	expression tag	UNP Q5SKW1
F	-9	SER	-	expression tag	UNP Q5SKW1
F	-8	SER	-	expression tag	UNP Q5SKW1
F	-7	GLY	-	expression tag	UNP Q5SKW1
F	-6	LEU	-	expression tag	UNP Q5SKW1
F	-5	VAL	-	expression tag	UNP Q5SKW1
F	-4	PRO	-	expression tag	UNP Q5SKW1
F	-3	ARG	-	expression tag	UNP Q5SKW1
F	-2	GLY	-	expression tag	UNP Q5SKW1
F	-1	SER	-	expression tag	UNP Q5SKW1
F	0	HIS	-	expression tag	UNP Q5SKW1

- Molecule 6 is a DNA chain called nontemplate DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
6	H	24	495	236	94	142	23	0	0	0

- Molecule 7 is a DNA chain called template DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
7	G	10	202	97	38	58	9	0	0	0

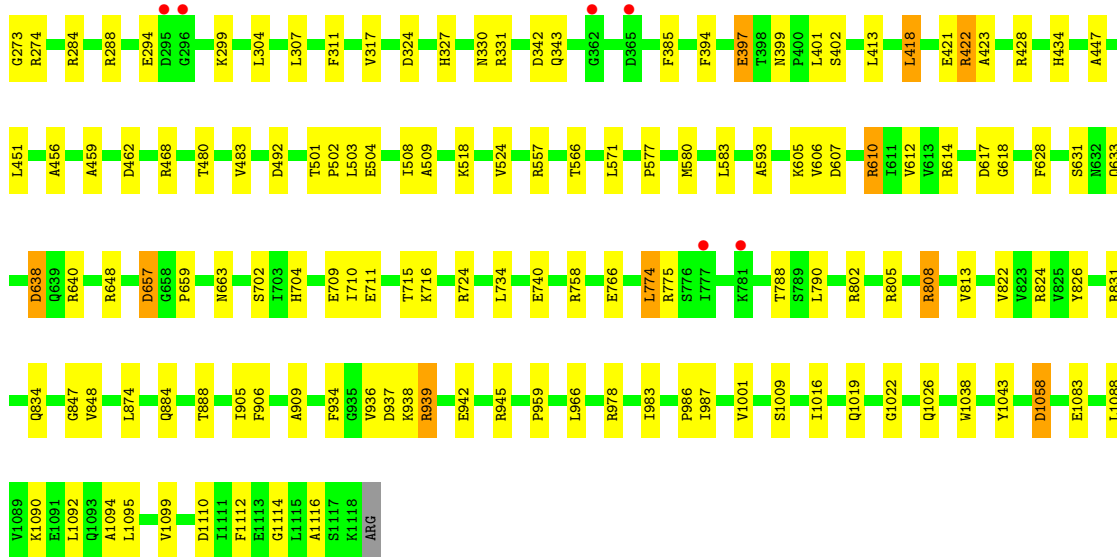
- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	B	1	Total 1	Mg 1	0	0
8	D	3	Total 3	Mg 3	0	0
8	F	1	Total 1	Mg 1	0	0

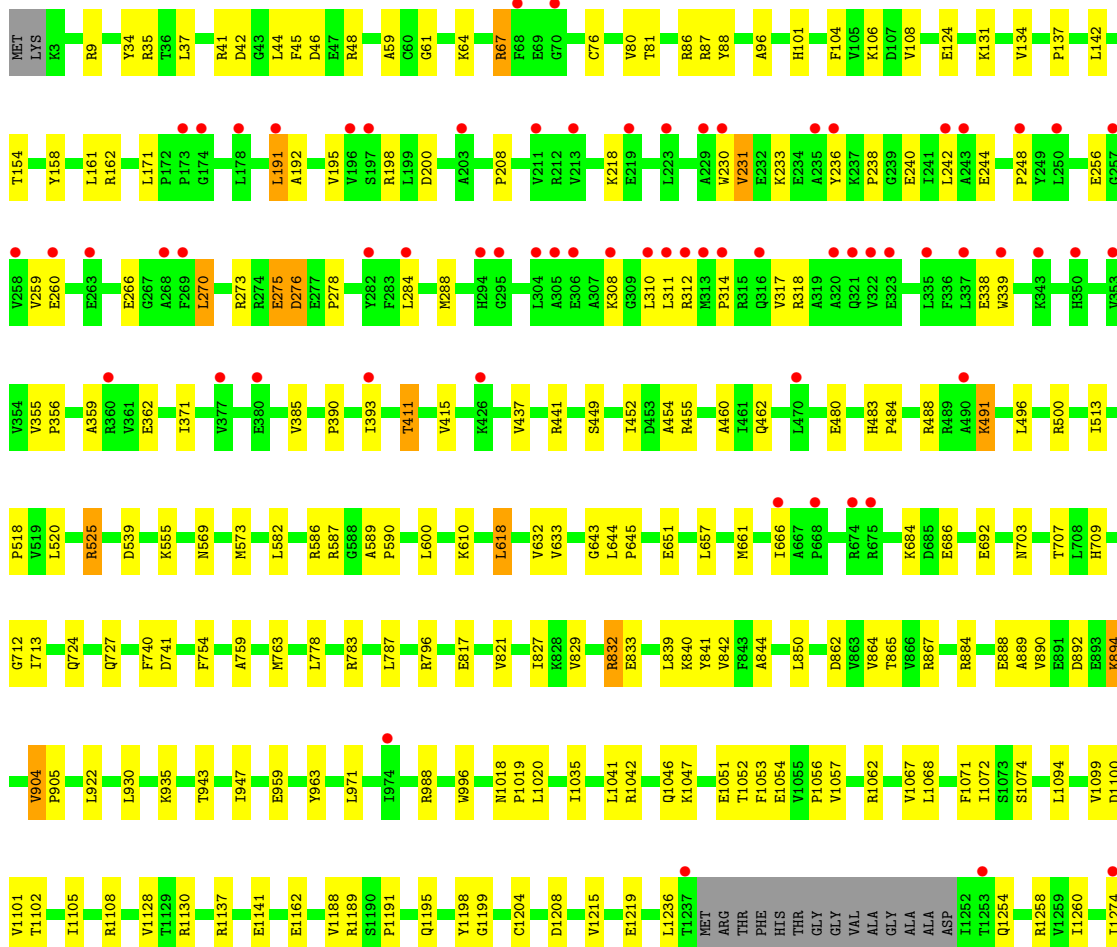
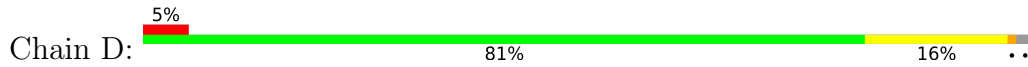
- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	2	Total 2	Zn 2	0	0





• Molecule 3: DNA-directed RNA polymerase subunit beta'







## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	183.70Å 103.85Å 295.41Å 90.00° 98.61° 90.00°	Depositor
Resolution (Å)	44.23 – 3.45 44.23 – 3.45	Depositor EDS
% Data completeness (in resolution range)	99.7 (44.23-3.45) 99.8 (44.23-3.45)	Depositor EDS
$R_{merge}$	0.17	Depositor
$R_{sym}$	0.17	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.18 (at 3.48Å)	Xtrriage
Refinement program	PHENIX 1.13_2998	Depositor
R, $R_{free}$	0.220 , 0.258 0.220 , 0.258	Depositor DCC
$R_{free}$ test set	1606 reflections (2.21%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	81.4	Xtrriage
Anisotropy	0.465	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 42.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	28227	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	101.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.02% 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: MG, ZN

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	0.22	0/1814	0.43	0/2466
1	B	0.22	0/1782	0.43	0/2424
2	C	0.22	0/8937	0.42	0/12087
3	D	0.22	0/11944	0.41	0/16149
4	E	0.21	0/772	0.38	0/1040
5	F	0.22	0/2766	0.38	0/3719
6	H	0.46	1/556 (0.2%)	0.95	1/858 (0.1%)
7	G	0.20	0/226	0.58	0/347
All	All	0.23	1/28797 (0.0%)	0.43	1/39090 (0.0%)

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
3	D	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	H	14	DG	O3'-P	6.02	1.68	1.61

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	H	14	DG	P-O3'-C3'	6.46	127.45	119.70

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	108	VAL	Peptide

## 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	1782	0	1834	26	0
1	B	1750	0	1797	26	0
2	C	8770	0	8874	106	0
3	D	11738	0	11972	140	0
4	E	758	0	770	8	0
5	F	2725	0	2810	31	0
6	H	495	0	272	17	0
7	G	202	0	114	5	0
8	B	1	0	0	0	0
8	D	3	0	0	0	0
8	F	1	0	0	0	0
9	D	2	0	0	0	0
All	All	28227	0	28443	318	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 6.

The worst 5 of 318 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)
6:H:22:DT:H2''	6:H:23:DG:H5'	1.58	0.83
2:C:802:ARG:HB2	2:C:826:TYR:HB2	1.67	0.77
2:C:628:PHE:H	2:C:638:ASP:HB2	1.51	0.76
2:C:165:LEU:HB2	2:C:168:ARG:HG3	1.69	0.75
1:B:206:THR:HG22	1:B:209:GLU:H	1.53	0.73

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	224/315 (71%)	221 (99%)	3 (1%)	0	100	100
1	B	220/315 (70%)	215 (98%)	5 (2%)	0	100	100
2	C	1107/1119 (99%)	1076 (97%)	31 (3%)	0	100	100
3	D	1482/1524 (97%)	1442 (97%)	40 (3%)	0	100	100
4	E	92/99 (93%)	90 (98%)	2 (2%)	0	100	100
5	F	332/443 (75%)	328 (99%)	4 (1%)	0	100	100
All	All	3457/3815 (91%)	3372 (98%)	85 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	199/273 (73%)	196 (98%)	3 (2%)	65	84
1	B	195/273 (71%)	189 (97%)	6 (3%)	40	70
2	C	936/941 (100%)	896 (96%)	40 (4%)	29	61
3	D	1253/1279 (98%)	1201 (96%)	52 (4%)	30	61
4	E	82/88 (93%)	81 (99%)	1 (1%)	71	87
5	F	292/388 (75%)	281 (96%)	11 (4%)	33	64
All	All	2957/3242 (91%)	2844 (96%)	113 (4%)	33	64

5 of 113 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	256	GLU
5	F	377	ASP
3	D	686	GLU
5	F	364	ARG
3	D	1496	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	1195	GLN
3	D	1333	HIS
5	F	214	GLN
3	D	1441	GLN
2	C	834	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 monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 7 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

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	226/315 (71%)	-0.01	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	69, 93, 116, 124	0
1	B	222/315 (70%)	0.09	1 (0%) <span style="border: 1px solid blue; padding: 2px;">91</span> <span style="border: 1px solid blue; padding: 2px;">89</span>	72, 104, 134, 151	0
2	C	1111/1119 (99%)	0.01	23 (2%) <span style="border: 1px solid blue; padding: 2px;">63</span> <span style="border: 1px solid blue; padding: 2px;">61</span>	57, 86, 156, 192	0
3	D	1486/1524 (97%)	0.22	79 (5%) <span style="border: 1px solid red; padding: 2px;">26</span> <span style="border: 1px solid red; padding: 2px;">26</span>	55, 89, 157, 186	1 (0%)
4	E	94/99 (94%)	0.02	2 (2%) <span style="border: 1px solid blue; padding: 2px;">63</span> <span style="border: 1px solid blue; padding: 2px;">61</span>	62, 92, 133, 142	0
5	F	336/443 (75%)	0.21	13 (3%) <span style="border: 1px solid red; padding: 2px;">39</span> <span style="border: 1px solid red; padding: 2px;">38</span>	67, 108, 165, 181	0
6	H	24/27 (88%)	0.63	4 (16%) <span style="border: 1px solid red; padding: 2px;">1</span> <span style="border: 1px solid red; padding: 2px;">2</span>	94, 158, 201, 220	0
7	G	10/19 (52%)	1.06	1 (10%) <span style="border: 1px solid red; padding: 2px;">7</span> <span style="border: 1px solid red; padding: 2px;">9</span>	149, 177, 201, 207	0
All	All	3509/3861 (90%)	0.13	123 (3%) <span style="border: 1px solid red; padding: 2px;">44</span> <span style="border: 1px solid red; padding: 2px;">42</span>	55, 93, 158, 220	1 (0%)

The worst 5 of 123 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	1237	THR	8.0
3	D	322	VAL	5.8
5	F	149	GLU	5.6
3	D	1297	GLU	5.3
2	C	219	GLN	4.6

### 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 monosaccharides 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(Å <sup>2</sup> )	Q<0.9
8	MG	B	2001	1/1	0.70	0.79	120,120,120,120	0
8	MG	D	2005	1/1	0.86	0.14	91,91,91,91	0
8	MG	F	2001	1/1	0.91	0.24	91,91,91,91	0
8	MG	D	2004	1/1	0.97	0.46	62,62,62,62	0
8	MG	D	2003	1/1	0.98	0.34	53,53,53,53	0
9	ZN	D	2001	1/1	0.99	0.20	60,60,60,60	0
9	ZN	D	2002	1/1	0.99	0.10	101,101,101,101	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.