



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 20, 2024 – 11:01 AM EST

PDB ID : 4LK0  
Title : Crystal Structure Analysis of the E.coli holoenzyme/T7 Gp2 complex  
Authors : Bae, B.; Darst, S.A.  
Deposited on : 2013-07-05  
Resolution : 3.91 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  
Xtrriage (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

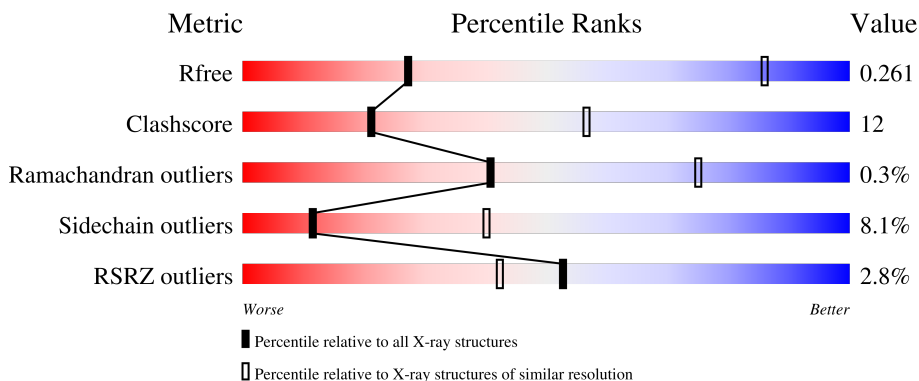
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.91 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	1019 (4.18-3.66)
Clashscore	141614	1016 (4.16-3.68)
Ramachandran outliers	138981	1039 (4.18-3.66)
Sidechain outliers	138945	1032 (4.18-3.66)
RSRZ outliers	127900	1002 (4.20-3.64)

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	239	 2% 60% 32% • 6%
1	B	239	 6% 64% 24% • 8%
1	G	239	 5% 61% 33% • 5%
1	H	239	 6% 60% 27% • 9%
2	C	1342	 3% 68% 29% •

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	I	1342	
3	D	1407	
3	J	1407	
4	E	91	
4	K	91	
5	F	522	
5	L	522	
6	M	64	
6	N	64	

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
7	MG	D	1501	-	-	-	X

## 2 Entry composition i

There are 8 unique types of molecules in this entry. The entry contains 58505 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	224	Total 1730	C 1076	N 308	O 340	S 6	0	0	0
1	B	220	Total 1687	C 1053	N 298	O 330	S 6	0	0	0
1	G	228	Total 1750	C 1088	N 312	O 344	S 6	0	0	0
1	H	217	Total 1667	C 1041	N 293	O 327	S 6	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	235	GLU	-	expression tag	UNP C9QXI7
A	236	VAL	-	expression tag	UNP C9QXI7
A	237	LEU	-	expression tag	UNP C9QXI7
A	238	PHE	-	expression tag	UNP C9QXI7
A	239	GLN	-	expression tag	UNP C9QXI7
B	235	GLU	-	expression tag	UNP C9QXI7
B	236	VAL	-	expression tag	UNP C9QXI7
B	237	LEU	-	expression tag	UNP C9QXI7
B	238	PHE	-	expression tag	UNP C9QXI7
B	239	GLN	-	expression tag	UNP C9QXI7
G	235	GLU	-	expression tag	UNP C9QXI7
G	236	VAL	-	expression tag	UNP C9QXI7
G	237	LEU	-	expression tag	UNP C9QXI7
G	238	PHE	-	expression tag	UNP C9QXI7
G	239	GLN	-	expression tag	UNP C9QXI7
H	235	GLU	-	expression tag	UNP C9QXI7
H	236	VAL	-	expression tag	UNP C9QXI7
H	237	LEU	-	expression tag	UNP C9QXI7
H	238	PHE	-	expression tag	UNP C9QXI7
H	239	GLN	-	expression tag	UNP C9QXI7

- 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	1340	Total	C	N	O	S	0	0	0
			10570	6631	1841	2055	43			
2	I	1340	Total	C	N	O	S	0	0	0
			10566	6629	1840	2054	43			

- 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	1345	Total	C	N	O	S	0	0	0
			10447	6560	1864	1974	49			
3	J	1325	Total	C	N	O	S	0	0	0
			10295	6470	1831	1945	49			

- 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	89	Total	C	N	O	S	0	0	0
			691	421	129	140	1			
4	K	79	Total	C	N	O	S	0	0	0
			627	382	118	126	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	F	470	Total	C	N	O	S	0	0	0
			3822	2394	680	725	23			
5	L	469	Total	C	N	O	S	0	0	0
			3821	2393	679	726	23			

- Molecule 6 is a protein called Bacterial RNA polymerase inhibitor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	M	51	Total	C	N	O	S	0	0	0
			413	268	65	79	1			
6	N	51	Total	C	N	O	S	0	0	0
			413	268	65	79	1			

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	D	1	Total Mg 1 1	0	0
7	J	1	Total Mg 1 1	0	0

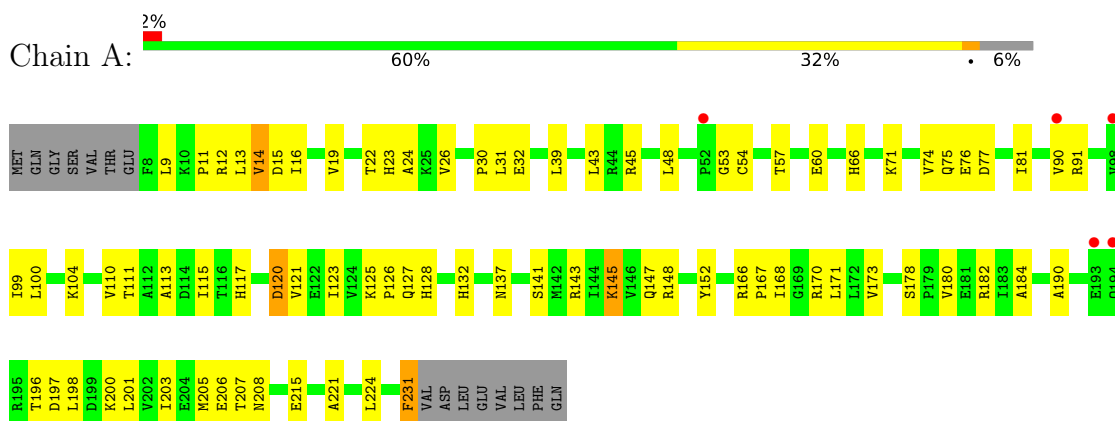
- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	D	2	Total Zn 2 2	0	0
8	J	2	Total Zn 2 2	0	0

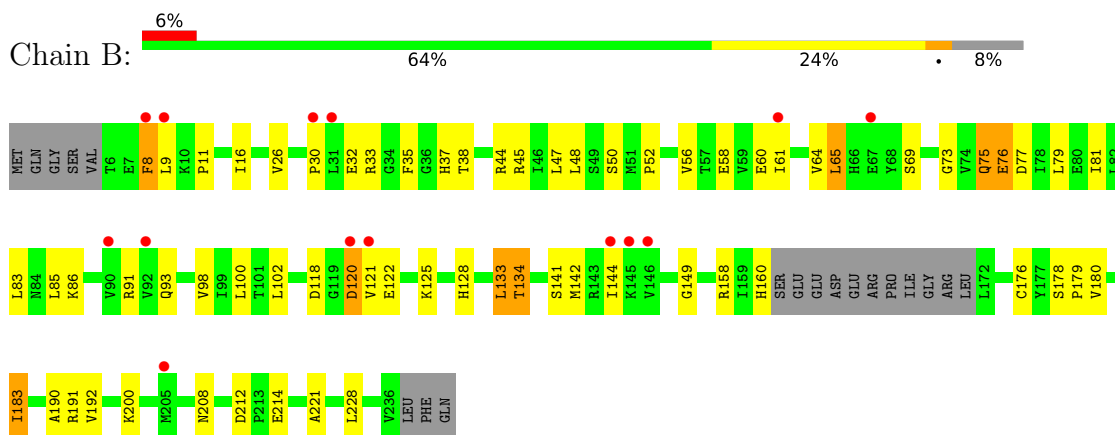
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

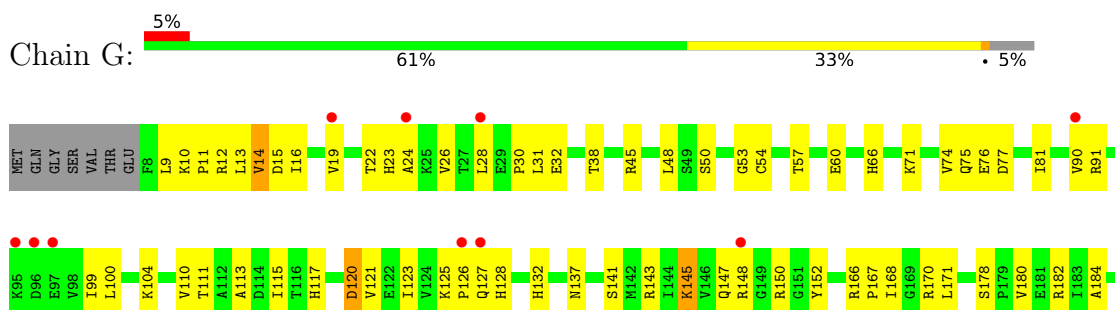
- Molecule 1: DNA-directed RNA polymerase subunit alpha

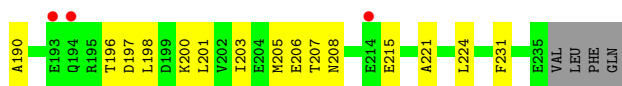


- Molecule 1: DNA-directed RNA polymerase subunit alpha

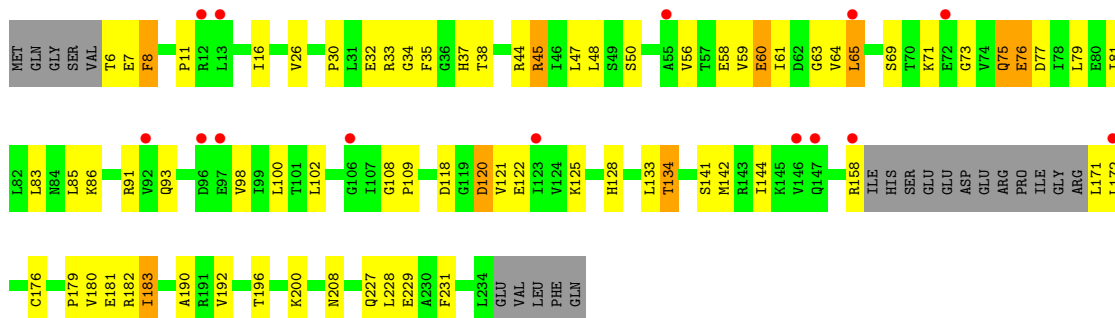


- Molecule 1: DNA-directed RNA polymerase subunit alpha

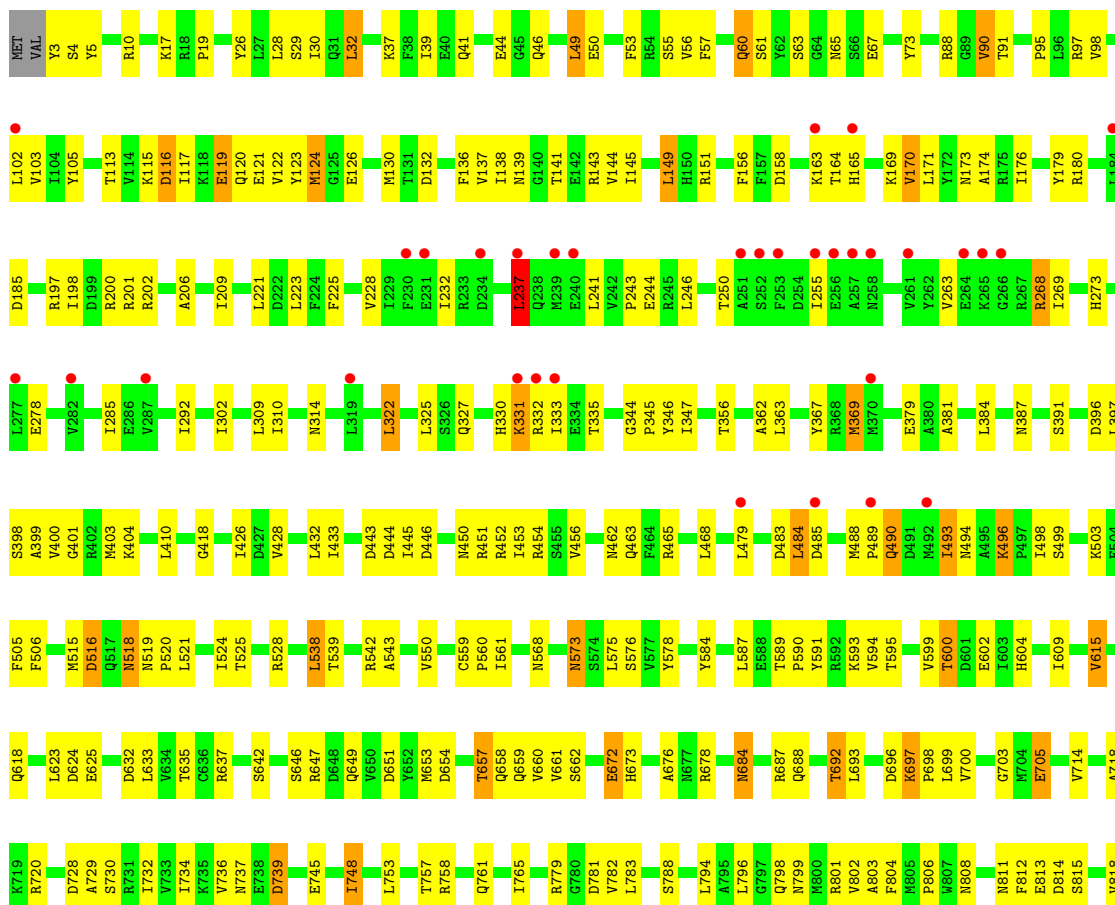




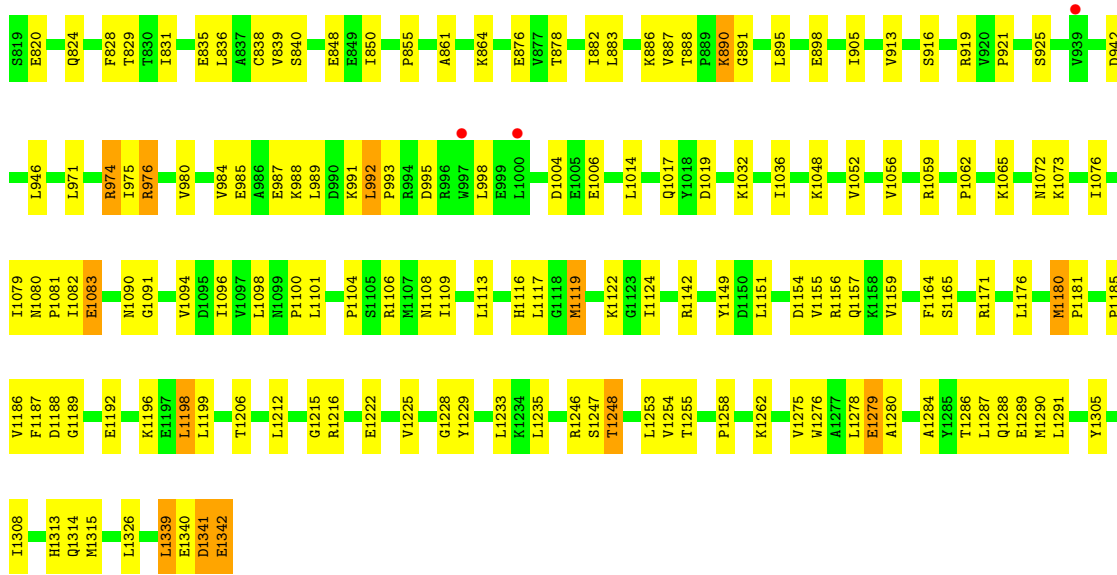
• Molecule 1: DNA-directed RNA polymerase subunit alpha



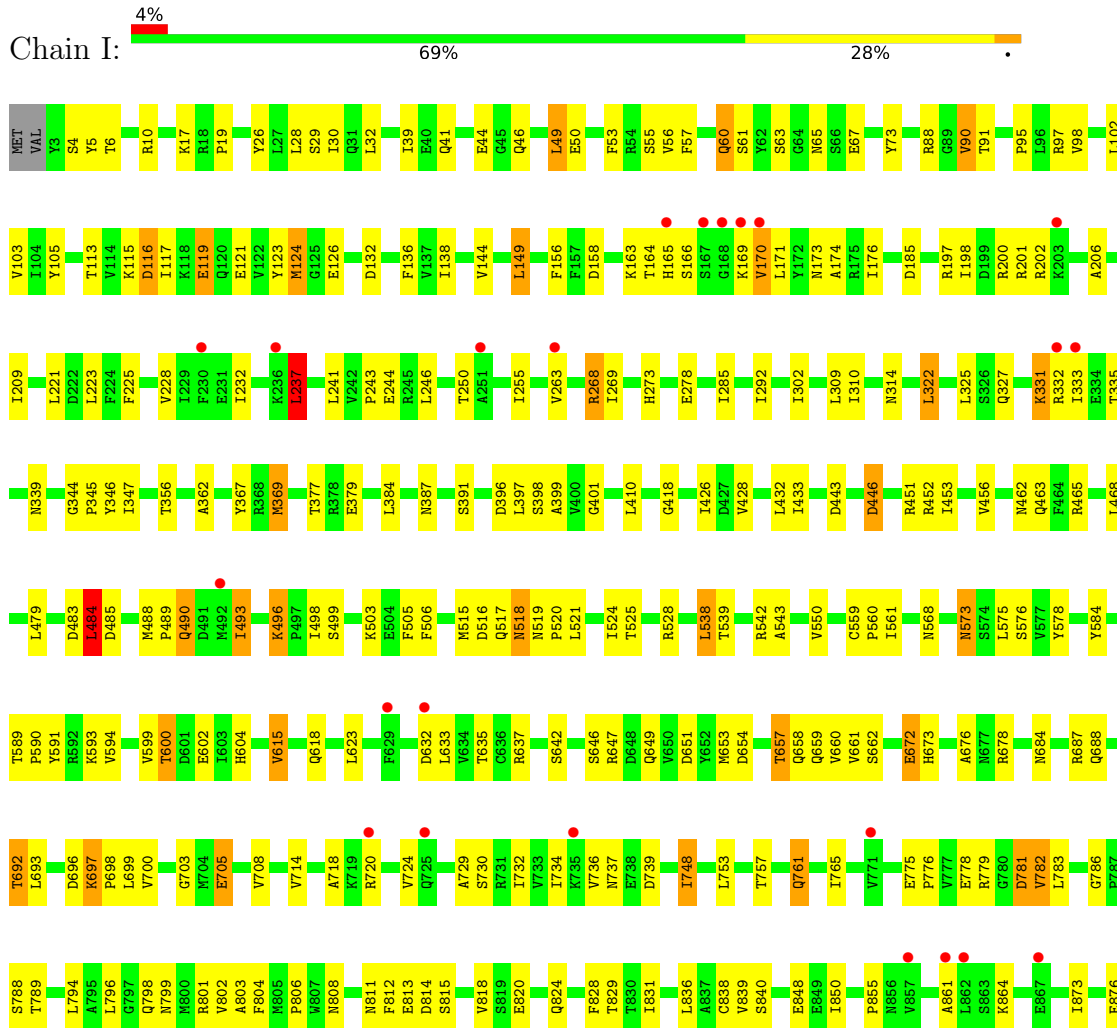
• Molecule 2: DNA-directed RNA polymerase subunit beta

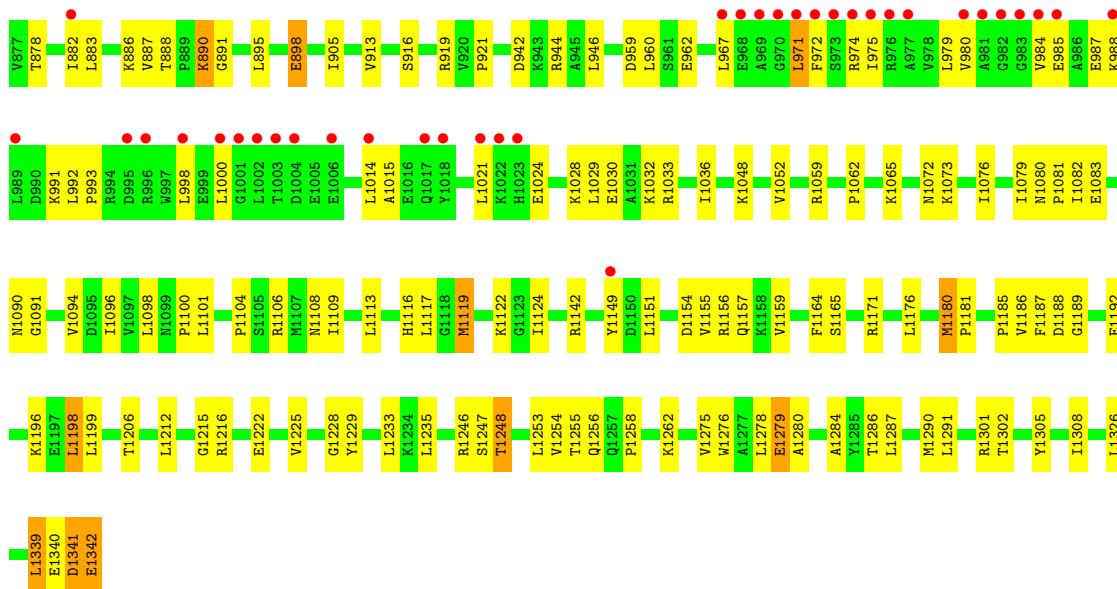




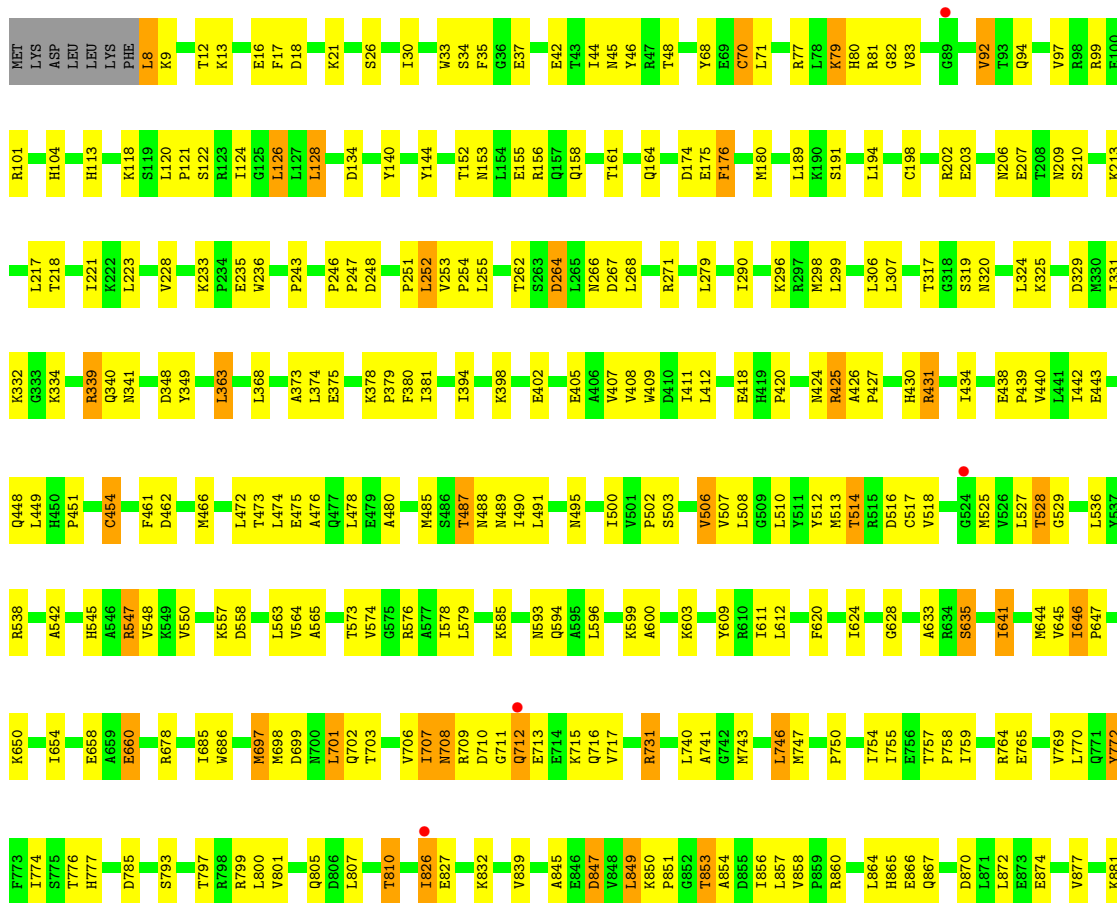


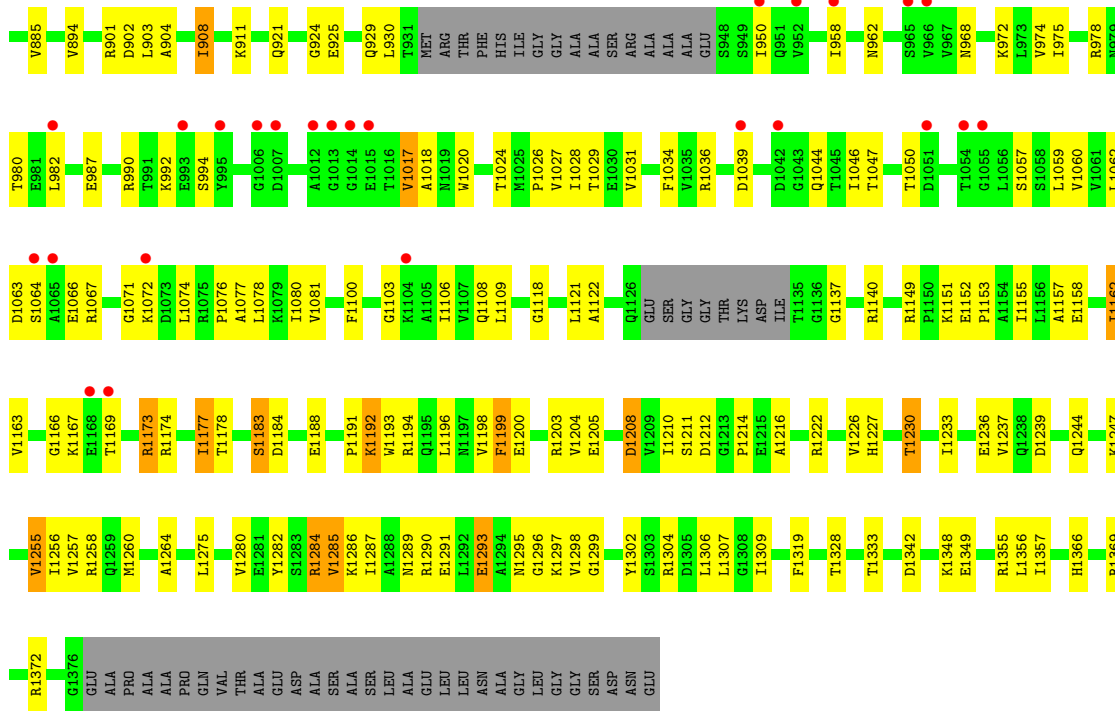
• Molecule 2: DNA-directed RNA polymerase subunit beta



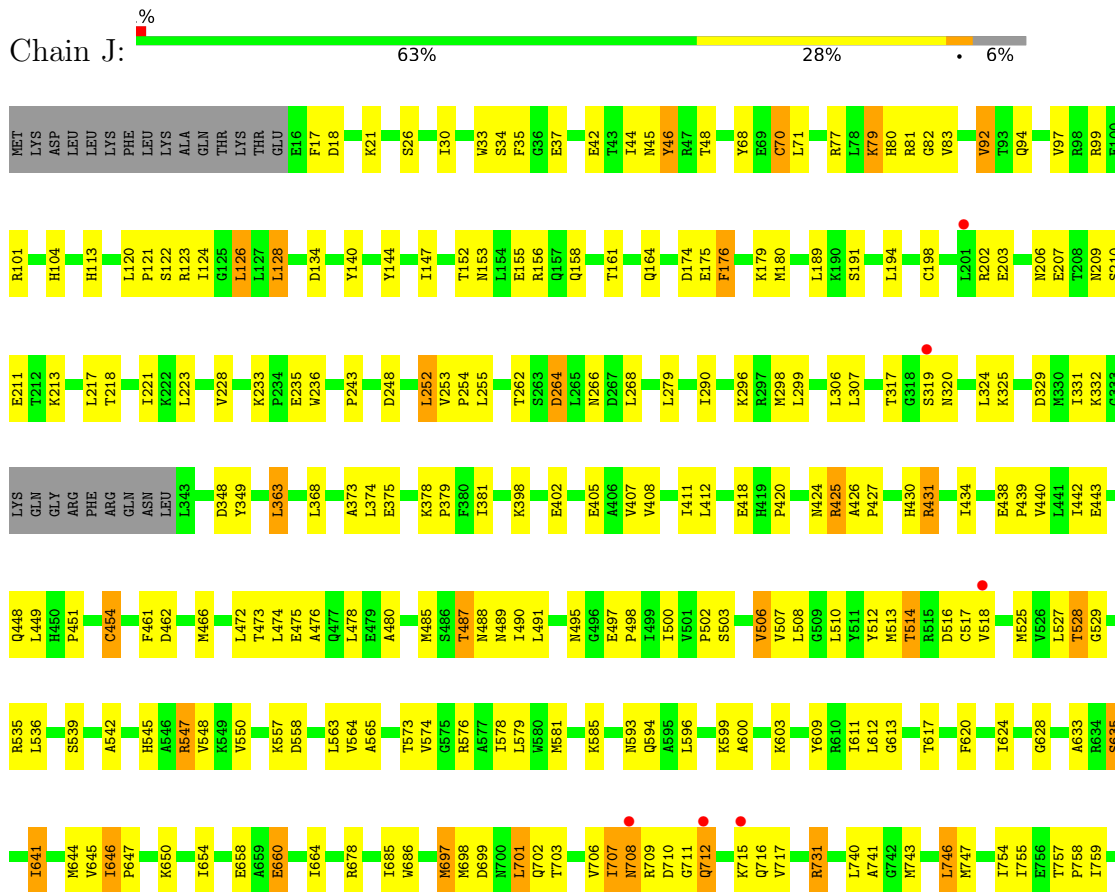


• Molecule 3: DNA-directed RNA polymerase subunit beta'



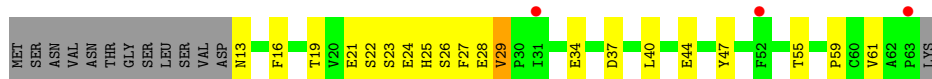


• Molecule 3: DNA-directed RNA polymerase subunit beta'









## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	186.36Å 206.28Å 308.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.94 – 3.91 39.94 – 3.91	Depositor EDS
% Data completeness (in resolution range)	92.8 (39.94-3.91) 92.9 (39.94-3.91)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.65 (at 3.87Å)	Xtrriage
Refinement program	PHENIX 1.8.1_1168	Depositor
R, $R_{free}$	0.219 , 0.260 0.220 , 0.261	Depositor DCC
$R_{free}$ test set	5023 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	120.8	Xtrriage
Anisotropy	0.647	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.25 , 47.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	58505	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	82.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.69% 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: ZN, 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	0.25	0/1751	0.51	0/2373
1	B	0.27	0/1707	0.51	0/2314
1	G	0.25	0/1771	0.52	0/2401
1	H	0.27	0/1686	0.51	0/2285
2	C	0.26	0/10739	0.49	2/14489 (0.0%)
2	I	0.26	0/10735	0.49	2/14484 (0.0%)
3	D	0.25	0/10603	0.48	0/14316
3	J	0.25	0/10450	0.47	0/14112
4	E	0.24	0/693	0.47	0/935
4	K	0.23	0/629	0.47	0/847
5	F	0.26	0/3873	0.49	1/5206 (0.0%)
5	L	0.27	0/3872	0.49	1/5205 (0.0%)
6	M	0.25	0/426	0.41	0/583
6	N	0.24	0/426	0.41	0/583
All	All	0.25	0/59361	0.48	6/80133 (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
3	J	0	1
All	All	0	2

There are no bond length outliers.

The worst 5 of 6 bond angle outliers are listed below:



Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	I	237	LEU	CA-CB-CG	6.80	130.94	115.30
2	C	237	LEU	CA-CB-CG	6.68	130.67	115.30
2	I	516	ASP	CB-CG-OD2	5.27	123.04	118.30
2	C	516	ASP	CB-CG-OD2	5.18	122.96	118.30
5	F	149	ASP	CB-CG-OD2	5.17	122.95	118.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	D	1296	GLY	Peptide
3	J	1296	GLY	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	1730	0	1756	53	0
1	B	1687	0	1700	43	0
1	G	1750	0	1764	57	0
1	H	1667	0	1689	49	0
2	C	10570	0	10582	269	0
2	I	10566	0	10576	248	0
3	D	10447	0	10672	304	0
3	J	10295	0	10511	292	0
4	E	691	0	695	15	0
4	K	627	0	634	12	0
5	F	3822	0	3885	101	0
5	L	3821	0	3884	105	0
6	M	413	0	389	10	0
6	N	413	0	389	13	0
7	D	1	0	0	0	0
7	J	1	0	0	0	0
8	D	2	0	0	0	0
8	J	2	0	0	0	0
All	All	58505	0	59126	1420	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 12.

The worst 5 of 1420 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:D:206:ASN:ND2	3:J:1183:SER:OG	2.04	0.90
3:D:1183:SER:OG	3:J:206:ASN:ND2	2.08	0.87
2:I:10:ARG:HD3	2:I:1181:PRO:HG2	1.60	0.83
2:C:452:ARG:NH1	2:C:584:TYR:O	2.11	0.83
2:C:10:ARG:HD3	2:C:1181:PRO:HG2	1.60	0.81

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	222/239 (93%)	197 (89%)	21 (10%)	4 (2%)	8	41
1	B	216/239 (90%)	189 (88%)	27 (12%)	0	100	100
1	G	226/239 (95%)	202 (89%)	20 (9%)	4 (2%)	8	41
1	H	213/239 (89%)	187 (88%)	26 (12%)	0	100	100
2	C	1338/1342 (100%)	1213 (91%)	121 (9%)	4 (0%)	41	75
2	I	1338/1342 (100%)	1216 (91%)	117 (9%)	5 (0%)	34	71
3	D	1339/1407 (95%)	1217 (91%)	120 (9%)	2 (0%)	51	83
3	J	1317/1407 (94%)	1201 (91%)	115 (9%)	1 (0%)	51	83
4	E	87/91 (96%)	79 (91%)	8 (9%)	0	100	100
4	K	77/91 (85%)	70 (91%)	7 (9%)	0	100	100
5	F	464/522 (89%)	415 (89%)	48 (10%)	1 (0%)	47	79
5	L	463/522 (89%)	415 (90%)	47 (10%)	1 (0%)	47	79
6	M	49/64 (77%)	42 (86%)	7 (14%)	0	100	100

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
6	N	49/64 (77%)	42 (86%)	7 (14%)	0	100	100
All	All	7398/7808 (95%)	6685 (90%)	691 (9%)	22 (0%)	41	75

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	170	VAL
2	I	170	VAL
3	D	340	GLN
1	G	167	PRO
2	I	484	LEU

### 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	191/206 (93%)	180 (94%)	11 (6%)	20	49
1	B	184/206 (89%)	167 (91%)	17 (9%)	9	34
1	G	191/206 (93%)	180 (94%)	11 (6%)	20	49
1	H	183/206 (89%)	166 (91%)	17 (9%)	9	33
2	C	1155/1157 (100%)	1052 (91%)	103 (9%)	9	36
2	I	1154/1157 (100%)	1056 (92%)	98 (8%)	10	39
3	D	1125/1168 (96%)	1036 (92%)	89 (8%)	12	41
3	J	1110/1168 (95%)	1023 (92%)	87 (8%)	12	41
4	E	72/75 (96%)	69 (96%)	3 (4%)	30	57
4	K	67/75 (89%)	65 (97%)	2 (3%)	41	64
5	F	417/462 (90%)	383 (92%)	34 (8%)	11	39
5	L	418/462 (90%)	386 (92%)	32 (8%)	13	41
6	M	44/56 (79%)	38 (86%)	6 (14%)	3	21
6	N	44/56 (79%)	38 (86%)	6 (14%)	3	21
All	All	6355/6660 (95%)	5839 (92%)	516 (8%)	11	40

5 of 516 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	J	1275	LEU
5	L	130	VAL
3	J	1258	ARG
3	D	908	ILE
3	D	847	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
5	F	446	GLN
2	I	686	GLN
5	L	446	GLN
3	J	206	ASN
3	D	206	ASN

### 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 6 ligands modelled in this entry, 6 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

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	224/239 (93%)	0.06	5 (2%) 62 52	48, 85, 138, 169	0
1	B	220/239 (92%)	0.37	14 (6%) 19 15	46, 110, 148, 166	0
1	G	228/239 (95%)	0.08	13 (5%) 23 20	48, 94, 141, 167	0
1	H	217/239 (90%)	0.33	14 (6%) 18 15	48, 111, 150, 167	0
2	C	1340/1342 (99%)	-0.08	36 (2%) 54 44	12, 74, 129, 163	0
2	I	1340/1342 (99%)	0.04	59 (4%) 34 29	12, 78, 143, 189	0
3	D	1345/1407 (95%)	-0.08	29 (2%) 62 52	13, 61, 142, 194	0
3	J	1325/1407 (94%)	-0.14	21 (1%) 72 62	10, 63, 139, 195	0
4	E	89/91 (97%)	-0.22	0 100 100	30, 77, 122, 138	0
4	K	79/91 (86%)	-0.28	0 100 100	32, 76, 121, 164	0
5	F	470/522 (90%)	-0.16	8 (1%) 70 61	25, 98, 147, 179	0
5	L	469/522 (89%)	-0.16	9 (1%) 66 58	28, 96, 150, 181	0
6	M	51/64 (79%)	0.53	1 (1%) 65 56	66, 115, 146, 158	0
6	N	51/64 (79%)	0.31	3 (5%) 22 18	64, 112, 143, 149	0
All	All	7448/7808 (95%)	-0.04	212 (2%) 53 42	10, 79, 142, 195	0

The worst 5 of 212 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	I	982	GLY	9.6
2	I	981	ALA	8.1
3	J	1059	LEU	6.5
2	I	995	ASP	5.6
1	B	146	VAL	5.5

## 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( $\text{\AA}^2$ )	Q<0.9
7	MG	D	1501	1/1	0.70	0.44	41,41,41,41	0
7	MG	J	1501	1/1	0.81	0.35	40,40,40,40	0
8	ZN	D	1502	1/1	0.90	0.18	193,193,193,193	0
8	ZN	J	1502	1/1	0.98	0.06	71,71,71,71	0
8	ZN	D	1503	1/1	0.99	0.21	36,36,36,36	0
8	ZN	J	1503	1/1	1.00	0.15	23,23,23,23	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.