



# Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 11:25 am BST

PDB ID : 6DT8  
Title : Bacteriophage N4 RNA polymerase II elongation complex 1  
Authors : Molodtsov, V.; Murakami, K.S.  
Deposited on : 2018-06-15  
Resolution : 3.20 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  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

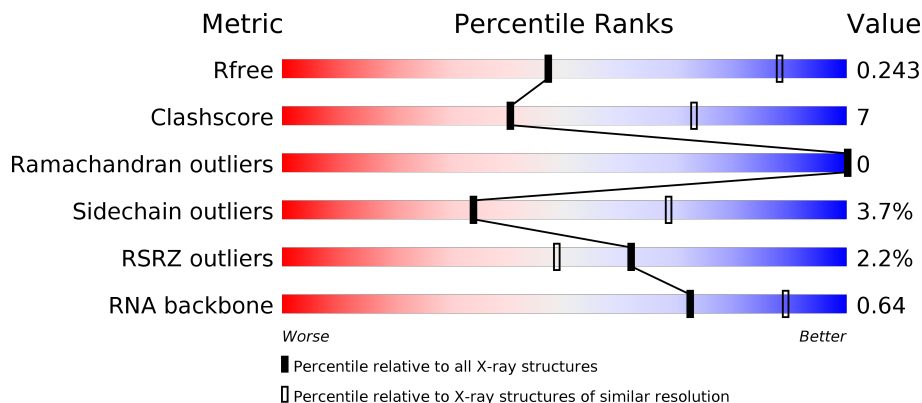
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashescore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)
RNA backbone	3102	1010 (3.50-2.90)

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 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	269	 78% 14% • 6%
2	B	404	 81% 16% •
3	C	49	 6% 10% 10% • 76%
4	D	12	 17% 17% 17% 50%

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 5672 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 RNAP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	252	2042	1295	352	382	13	0	0	0

- Molecule 2 is a protein called RNAP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	404	3256	2073	545	617	21	0	0	0

- Molecule 3 is a DNA chain called DNA (5'-D(P\*AP\*CP\*CP\*CP\*AP\*CP\*CP\*AP\*AP\*AP\*AP\*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	C	12	242	115	50	65	12	0	0	0

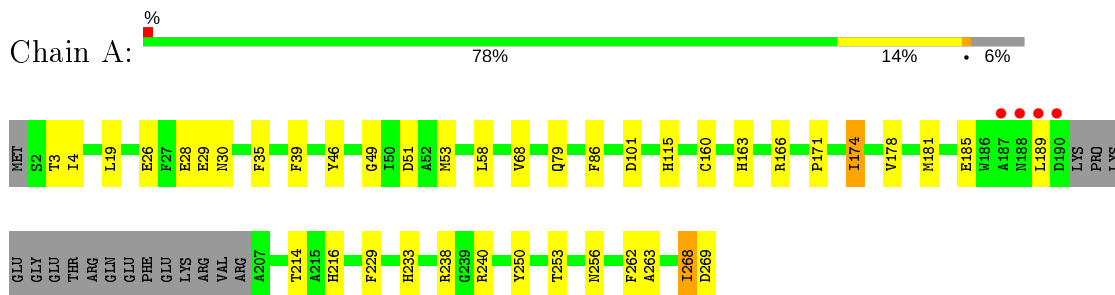
- Molecule 4 is a RNA chain called RNA (5'-R(P\*UP\*GP\*GP\*UP\*GP\*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
4	D	6	132	58	24	44	6	0	0	0

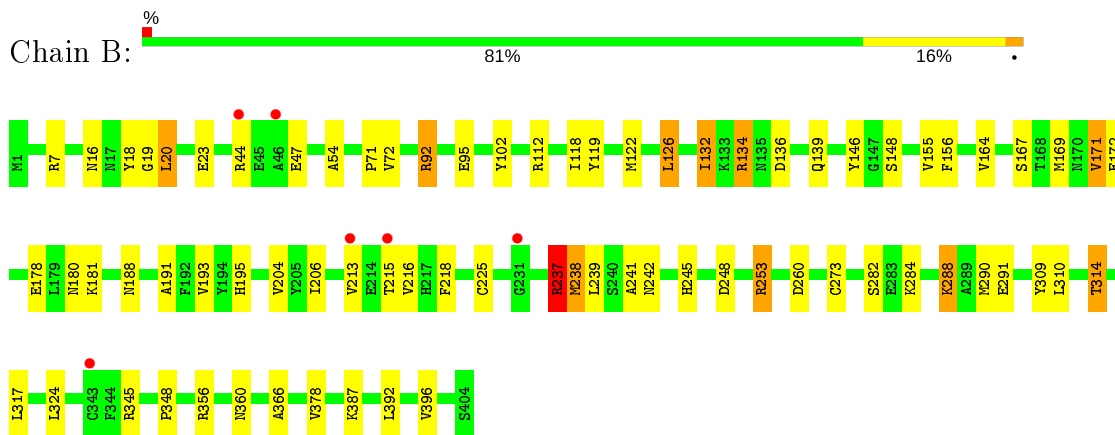
### 3 Residue-property plots [\*i\*](#)

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.

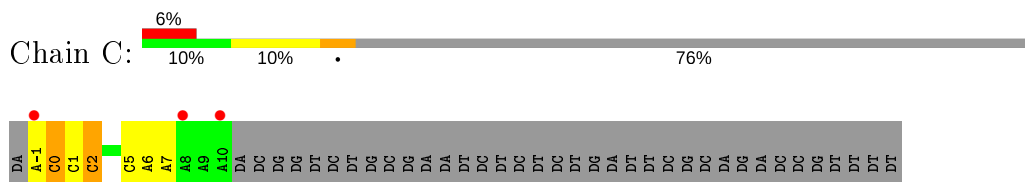
- Molecule 1: RNAP1



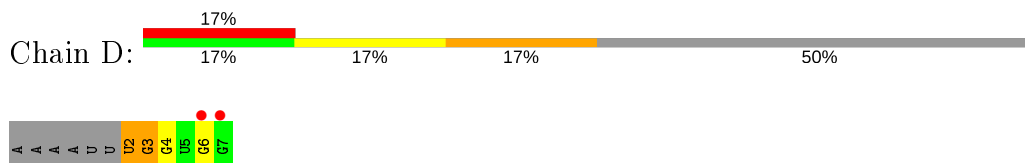
- Molecule 2: RNAP2



- Molecule 3: DNA (5'-D(P\*AP\*CP\*CP\*CP\*AP\*CP\*CP\*AP\*AP\*AP\*AP\*A)-3')



- Molecule 4: RNA (5'-R(P\*UP\*GP\*GP\*UP\*GP\*G)-3')



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 4	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	174.76Å 174.76Å 76.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.19 – 3.20 41.19 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.9 (41.19-3.20) 99.9 (41.19-3.20)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.76 (at 3.18Å)	Xtrriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, $R_{free}$	0.190 , 0.243 0.190 , 0.243	Depositor DCC
$R_{free}$ test set	1924 reflections (10.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	107.8	Xtrriage
Anisotropy	0.066	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 46.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.032 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	5672	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	106.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% 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](#)

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	0.55	1/2089 (0.0%)	0.77	0/2821
2	B	0.50	0/3323	0.77	4/4496 (0.1%)
3	C	0.88	0/272	1.05	2/415 (0.5%)
4	D	0.74	0/147	1.20	0/228
All	All	0.55	1/5831 (0.0%)	0.80	6/7960 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	160	CYS	CB-SG	-5.37	1.73	1.81

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	2	DC	O4'-C4'-C3'	-8.28	101.03	106.00
2	B	237	ARG	NE-CZ-NH1	7.56	124.08	120.30
2	B	237	ARG	CA-CB-CG	6.06	126.73	113.40
2	B	253	ARG	NE-CZ-NH1	5.70	123.15	120.30
2	B	237	ARG	CG-CD-NE	5.45	123.23	111.80
3	C	0	DC	O4'-C1'-N1	5.03	111.52	108.00

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	2042	0	1964	27	0
2	B	3256	0	3214	56	0
3	C	242	0	133	8	0
4	D	132	0	65	2	0
All	All	5672	0	5376	78	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 7.

All (78) 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:237:ARG:HH11	2:B:237:ARG:HB3	1.39	0.86
3:C:5:DC:H2''	3:C:6:DA:H5'	1.59	0.84
2:B:237:ARG:NH1	2:B:237:ARG:HB3	2.05	0.71
2:B:136:ASP:HB3	2:B:155:VAL:HG13	1.73	0.69
2:B:238:MET:HG3	2:B:242:ASN:ND2	2.08	0.68
1:A:115:HIS:CE1	1:A:214:THR:HG22	2.29	0.68
2:B:290:MET:HG2	2:B:317:LEU:HD22	1.76	0.67
1:A:4:ILE:HD12	1:A:4:ILE:H	1.61	0.66
1:A:253:THR:OG1	1:A:256:ASN:HB2	1.97	0.64
1:A:181:MET:O	1:A:185:GLU:HG2	1.99	0.62
1:A:19:LEU:HD21	1:A:58:LEU:HD21	1.81	0.62
2:B:237:ARG:HH21	3:C:-1:DA:H1'	1.64	0.62
2:B:310:LEU:HA	2:B:314:THR:CG2	2.31	0.61
4:D:2:U:O2'	4:D:3:G:OP1	2.18	0.61
1:A:68:VAL:HG21	2:B:225:CYS:SG	2.40	0.60
2:B:119:TYR:CD1	2:B:134:ARG:HA	2.37	0.60
3:C:6:DA:H2''	3:C:7:DA:O4'	2.01	0.59
2:B:148:SER:HA	3:C:0:DC:H5'	1.86	0.57
2:B:178:GLU:O	2:B:181:LYS:HB3	2.03	0.57
2:B:95:GLU:O	2:B:112:ARG:HD3	2.05	0.55
2:B:238:MET:HG3	2:B:242:ASN:HD22	1.72	0.55
1:A:269:ASP:OXT	2:B:7:ARG:NH1	2.40	0.55
2:B:132:ILE:HD12	2:B:156:PHE:HE1	1.73	0.54
2:B:239:LEU:HD13	2:B:378:VAL:HG11	1.90	0.54
2:B:366:ALA:O	2:B:387:LYS:HE2	2.07	0.53
2:B:72:VAL:O	2:B:345:ARG:HA	2.07	0.53
2:B:118:ILE:HD13	2:B:169:MET:SD	2.49	0.53
2:B:237:ARG:CB	2:B:237:ARG:HH11	2.17	0.52
2:B:309:TYR:O	2:B:314:THR:HG21	2.10	0.51
2:B:310:LEU:HA	2:B:314:THR:HG22	1.91	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:28:GLU:HG2	1:A:46:TYR:OH	2.11	0.51
2:B:310:LEU:HA	2:B:314:THR:HG21	1.93	0.50
2:B:392:LEU:O	2:B:396:VAL:HG23	2.12	0.50
1:A:35:PHE:HB3	1:A:39:PHE:CE2	2.47	0.49
2:B:18:TYR:O	2:B:47:GLU:N	2.32	0.49
1:A:268:ILE:CD1	2:B:348:PRO:HG2	2.43	0.49
1:A:26:GLU:HA	1:A:30:ASN:HB2	1.95	0.49
4:D:2:U:H2'	4:D:3:G:C8	2.48	0.49
1:A:79:GLN:HA	2:B:216:VAL:HG22	1.94	0.48
2:B:282:SER:OG	2:B:284:LYS:HG2	2.14	0.48
3:C:5:DC:H2'	3:C:6:DA:C8	2.47	0.48
2:B:122:MET:O	2:B:126:LEU:HB2	2.13	0.48
1:A:49:GLY:O	1:A:53:MET:HG2	2.14	0.47
2:B:195:HIS:HA	2:B:204:VAL:O	2.14	0.47
2:B:167:SER:O	2:B:171:VAL:HG12	2.15	0.47
2:B:20:LEU:HB3	2:B:23:GLU:HG3	1.97	0.47
2:B:238:MET:HG2	2:B:239:LEU:N	2.16	0.46
1:A:268:ILE:O	1:A:269:ASP:HB2	2.16	0.46
2:B:237:ARG:CD	3:C:0:DC:H5''	2.45	0.46
2:B:102:TYR:CB	2:B:112:ARG:HD2	2.46	0.46
1:A:181:MET:HG2	1:A:185:GLU:OE2	2.16	0.46
2:B:136:ASP:HB3	2:B:155:VAL:CG1	2.44	0.46
2:B:288:LYS:HE3	2:B:291:GLU:OE1	2.15	0.46
3:C:6:DA:OP2	3:C:6:DA:H3'	2.16	0.45
1:A:86:PHE:HB2	2:B:218:PHE:CE2	2.51	0.45
1:A:262:PHE:CE2	2:B:71:PRO:HB3	2.52	0.45
1:A:268:ILE:HA	1:A:268:ILE:HD13	1.75	0.45
1:A:238:ARG:NH1	2:B:245:HIS:CD2	2.85	0.45
1:A:171:PRO:HB2	1:A:263:ALA:HB3	1.99	0.44
1:A:240:ARG:NH1	2:B:248:ASP:OD1	2.51	0.44
2:B:188:ASN:ND2	2:B:191:ALA:HB2	2.33	0.44
1:A:189:LEU:CD2	2:B:44:ARG:HA	2.47	0.44
2:B:126:LEU:HD12	2:B:164:VAL:HG11	2.00	0.44
3:C:1:DC:H2'	3:C:2:DC:O4'	2.18	0.44
1:A:178:VAL:HG21	2:B:54:ALA:HB2	2.00	0.44
2:B:16:ASN:O	2:B:19:GLY:HA3	2.17	0.43
2:B:314:THR:O	2:B:317:LEU:HG	2.19	0.43
2:B:134:ARG:C	2:B:134:ARG:HD3	2.40	0.42
2:B:193:VAL:HA	2:B:206:ILE:O	2.19	0.42
2:B:356:ARG:NH1	2:B:360:ASN:OD1	2.41	0.41
1:A:174:ILE:HD12	1:A:229:PHE:CE2	2.56	0.41

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:273:CYS:HB2	2:B:324:LEU:HD21	2.03	0.41
2:B:146:TYR:HD1	2:B:241:ALA:HA	1.86	0.41
2:B:92:ARG:NH1	2:B:172:GLU:OE1	2.54	0.41
1:A:166:ARG:HD3	2:B:260:ASP:HB3	2.03	0.41
2:B:18:TYR:N	2:B:19:GLY:HA2	2.36	0.41
1:A:233:HIS:NE2	1:A:250:TYR:OH	2.43	0.41
1:A:163:HIS:CD2	2:B:253:ARG:HA	2.56	0.40

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	248/269 (92%)	240 (97%)	8 (3%)	0	100	100
2	B	402/404 (100%)	391 (97%)	11 (3%)	0	100	100
All	All	650/673 (97%)	631 (97%)	19 (3%)	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	222/243 (91%)	215 (97%)	7 (3%)	39	71

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	B	351/351 (100%)	337 (96%)	14 (4%)	31	66
All	All	573/594 (96%)	552 (96%)	21 (4%)	34	68

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	THR
1	A	29	GLU
1	A	51	ASP
1	A	101	ASP
1	A	174	ILE
1	A	216	HIS
1	A	268	ILE
2	B	20	LEU
2	B	92	ARG
2	B	126	LEU
2	B	132	ILE
2	B	134	ARG
2	B	139	GLN
2	B	171	VAL
2	B	180	ASN
2	B	213	VAL
2	B	215	THR
2	B	237	ARG
2	B	238	MET
2	B	288	LYS
2	B	314	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	115	HIS

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	D	6/12 (50%)	3 (50%)	1 (16%)

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	D	3	G
4	D	4	G
4	D	6	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
4	D	2	U

#### 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](#)

There are no ligands in this entry.

#### 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	252/269 (93%)	-0.06	4 (1%) 72 59	62, 104, 164, 214	0
2	B	404/404 (100%)	0.01	6 (1%) 73 61	58, 97, 153, 187	0
3	C	12/49 (24%)	0.77	3 (25%) 0 0	105, 150, 241, 242	0
4	D	6/12 (50%)	2.19	2 (33%) 0 0	72, 95, 122, 152	6 (100%)
All	All	674/734 (91%)	0.01	15 (2%) 62 48	58, 99, 159, 242	6 (0%)

All (15) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	6	G	4.7
4	D	7	G	4.6
1	A	190	ASP	4.6
1	A	189	LEU	4.2
1	A	188	ASN	4.1
1	A	187	ALA	3.1
3	C	10	DA	2.8
2	B	343	CYS	2.8
3	C	-1	DA	2.6
2	B	46	ALA	2.5
2	B	215	THR	2.3
2	B	44	ARG	2.2
2	B	231	GLY	2.1
3	C	8	DA	2.0
2	B	213	VAL	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](#)

There are no ligands in this entry.

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

There are no such residues in this entry.