



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 11, 2019 – 03:36 PM EST

PDB ID : 4V8S  
Title : Archaeal RNAP-DNA binary complex at 4.32Å  
Authors : Wojtas, M.N.; Mogni, M.; Millet, O.; Bell, S.D.; Abrescia, N.G.A.  
Deposited on : 2012-07-12  
Resolution : 4.32 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.0 (224370), CSD as540be (2019)  
Xtriage (Phenix) : 1.13  
EDS : 2.4  
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) : 2.4

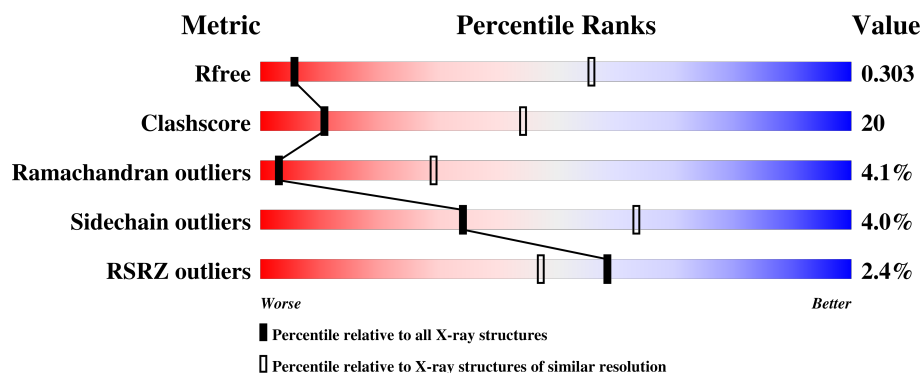
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.32 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.






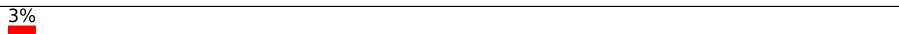
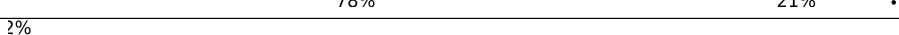
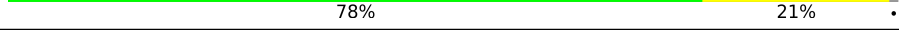



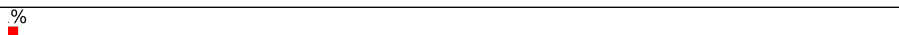
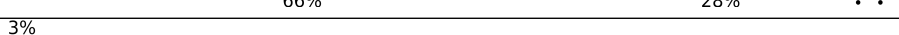




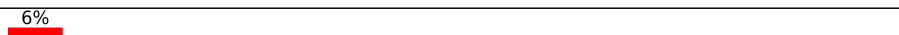


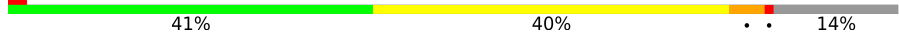
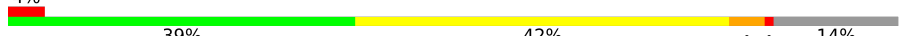


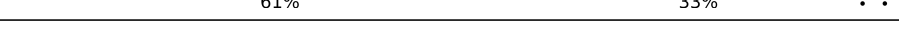


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	111664	1034 (4.92-3.70)
Clashscore	122126	1104 (4.92-3.70)
Ramachandran outliers	120053	1053 (4.92-3.70)
Sidechain outliers	120020	1037 (4.92-3.70)
RSRZ outliers	108989	1197 (5.04-3.60)

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	AC	14	<div> <div>21%</div> <div>57%</div> <div>29%</div> <div>7%</div> <div>7%</div> </div>
1	BR	14	<div> <div>14%</div> <div>50%</div> <div>43%</div> <div>7%</div> </div>
2	AD	16	<div> <div>13%</div> <div>19%</div> <div>63%</div> <div>13%</div> <div>6%</div> </div>
2	BS	16	<div> <div>13%</div> <div>19%</div> <div>75%</div> <div>6%</div> </div>
3	AI	95	<div> <div>62%</div> <div>23%</div> <div>•</div> <div>12%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	BK	95	
4	AJ	104	
4	BQ	104	
5	AM	92	
5	BL	92	
6	AO	66	
6	BN	66	
7	AR	1131	
7	BB	1131	
8	AS	265	
8	BD	265	
9	AT	180	
9	BE	180	
10	AU	113	
10	BF	113	
11	AV	132	
11	BG	132	
12	AW	880	
12	BA	880	
13	AX	48	
13	BP	48	
14	AY	395	
14	BC	395	
15	AZ	84	
15	BH	84	

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
17	SF4	AS	1001	-	-	X	-
17	SF4	BD	1001	-	-	X	-

## 2 Entry composition

There are 18 unique types of molecules in this entry. The entry contains 111598 atoms, of which 56187 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 DNA chain called 5'-D(\*TP\*CP\*TP\*TP\*AP\*TP\*AP\*CP\*TP\*CP\*TP\*AP\*TP\*CP)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AC	13	Total	C	H	N	O	P	0	0	0
			411	127	151	38	82	13			
1	BR	14	Total	C	H	N	O	P	0	0	0
			441	136	162	41	88	14			

- Molecule 2 is a DNA chain called 5'-D(\*AP\*TP\*AP\*GP\*AP\*GP\*TP\*AP\*TP\*AP\*AP\*GP\*AP\*TP\*AP\*G)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	AD	15	Total	C	H	N	O	P	0	0	0
			485	150	170	63	87	15			
2	BS	16	Total	C	H	N	O	P	0	0	0
			517	160	181	68	92	16			

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT K.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	AI	84	Total	C	H	N	O	S	0	0	0
			1390	431	717	123	118	1			
3	BK	84	Total	C	H	N	O	S	0	0	0
			1390	431	717	123	118	1			

- Molecule 4 is a protein called RNA POLYMERASE SUBUNIT 13.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
4	AJ	49	Total	C	H	N	O	S	0	0	0
			830	264	413	70	82	1			
4	BQ	50	Total	C	H	N	O	S	0	0	0
			845	269	419	71	85	1			

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT L.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
5	AM	91	Total	C	H	N	O	S	0	0	0
			1449	454	742	114	137	2			
5	BL	91	Total	C	H	N	O	S	0	0	0
			1449	454	742	114	137	2			

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT N.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
6	AO	65	Total	C	H	N	O	S	0	0	0
			1058	332	537	94	88	7			
6	BN	65	Total	C	H	N	O	S	0	0	0
			1058	332	537	94	88	7			

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
7	AR	1103	Total	C	H	N	O	S	0	0	0
			17665	5548	8909	1552	1627	29			
7	BB	1103	Total	C	H	N	O	S	0	0	0
			17665	5548	8909	1552	1627	29			

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT D.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
8	AS	262	Total	C	H	N	O	S	0	0	0
			4215	1339	2128	337	398	13			
8	BD	262	Total	C	H	N	O	S	0	0	0
			4215	1339	2128	337	398	13			

- Molecule 9 is a protein called RNA POLYMERASE SUBUNIT 7.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
9	AT	171	Total	C	H	N	O	S	0	0	0
			2772	874	1413	229	251	5			
9	BE	171	Total	C	H	N	O	S	0	0	0
			2771	874	1412	229	251	5			

- Molecule 10 is a protein called RNA POLYMERASE SUBUNIT 4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
10	AU	105	Total	C	H	N	O	S	0	0	0
			1667	519	840	134	171	3			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
10	BF	105	Total	C	H	N	O	S	0	0	0
			1667	519	840	134	171	3			

- Molecule 11 is a protein called RNA POLYMERASE SUBUNIT 8.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
11	AV	113	Total	C	H	N	O	S	0	0	0
			1816	572	915	152	173	4			
11	BG	113	Total	C	H	N	O	S	0	0	0
			1816	572	915	152	173	4			

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
12	AW	872	Total	C	H	N	O	S	0	0	0
			13987	4424	7030	1225	1282	26			
12	BA	872	Total	C	H	N	O	S	0	0	0
			13987	4424	7030	1225	1282	26			

- Molecule 13 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT P.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
13	AX	44	Total	C	H	N	O	S	0	0	0
			744	236	387	62	54	5			
13	BP	44	Total	C	H	N	O	S	0	0	0
			744	236	387	62	54	5			

- Molecule 14 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT A'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
14	AY	376	Total	C	H	N	O	S	0	0	0
			5974	1840	3068	493	564	9			
14	BC	376	Total	C	H	N	O	S	0	0	0
			5974	1840	3068	493	564	9			

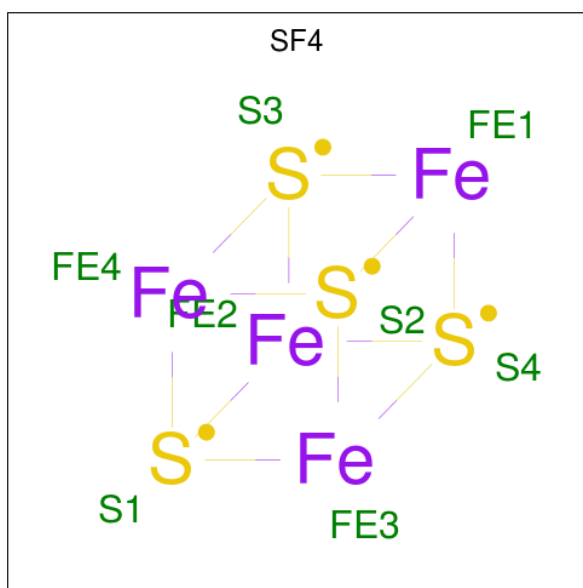
- Molecule 15 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT H.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
15	AZ	76	Total	C	H	N	O		0	0	0
			1284	405	660	111	108				
15	BH	76	Total	C	H	N	O		0	0	0
			1284	405	660	111	108				

- Molecule 16 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	BB	1	Total	Zn	0	0
			1	1		
16	BA	3	Total	Zn	0	0
			3	3		
16	BN	1	Total	Zn	0	0
			1	1		
16	AW	3	Total	Zn	0	0
			3	3		
16	BP	1	Total	Zn	0	0
			1	1		
16	AX	1	Total	Zn	0	0
			1	1		
16	AO	1	Total	Zn	0	0
			1	1		
16	AR	1	Total	Zn	0	0
			1	1		

- Molecule 17 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe<sub>4</sub>S<sub>4</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
17	AS	1	Total	Fe	S	0	0
			7	3	4		
17	BD	1	Total	Fe	S	0	0
			7	3	4		

- Molecule 18 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).



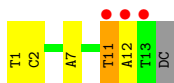
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
18	AW	1	Total 1	Mg 1	0	0
18	BA	1	Total 1	Mg 1	0	0

### 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: 5'-D(\*TP\*CP\*TP\*TP\*AP\*TP\*AP\*CP\*TP\*CP\*TP\*AP\*TP\*CP)-3'

Chain AC: 



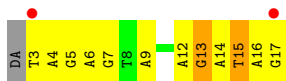
- Molecule 1: 5'-D(\*TP\*CP\*TP\*TP\*AP\*TP\*AP\*CP\*TP\*CP\*TP\*AP\*TP\*CP)-3'

Chain BR: 



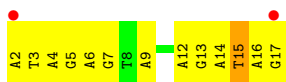
- Molecule 2: 5'-D(\*AP\*TP\*AP\*GP\*AP\*GP\*TP\*AP\*TP\*AP\*AP\*GP\*AP\*TP \*AP\*G)-3'

Chain AD: 



- Molecule 2: 5'-D(\*AP\*TP\*AP\*GP\*AP\*GP\*TP\*AP\*TP\*AP\*AP\*GP\*AP\*TP \*AP\*G)-3'

Chain BS: 



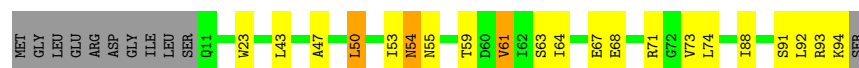
- Molecule 3: DNA-DIRECTED RNA POLYMERASE SUBUNIT K

Chain AI: 




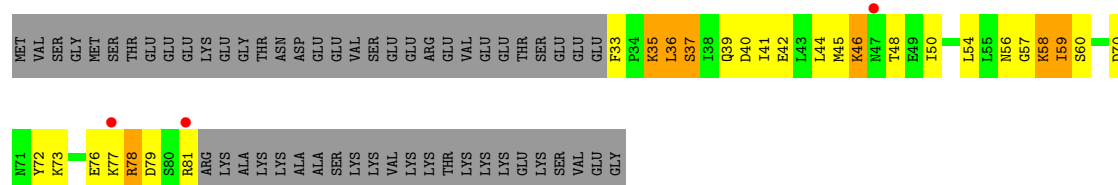
- Molecule 3: DNA-DIRECTED RNA POLYMERASE SUBUNIT K

Chain BK: 




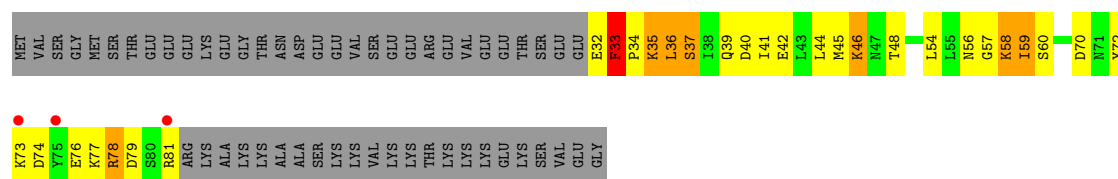
• Molecule 4: RNA POLYMERASE SUBUNIT 13

Chain AJ: 




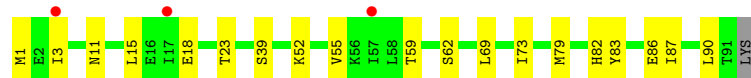
• Molecule 4: RNA POLYMERASE SUBUNIT 13

Chain BQ: 




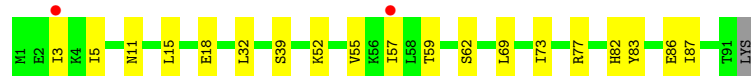
• Molecule 5: DNA-DIRECTED RNA POLYMERASE SUBUNIT L

Chain AM: 



• Molecule 5: DNA-DIRECTED RNA POLYMERASE SUBUNIT L

Chain BL: 



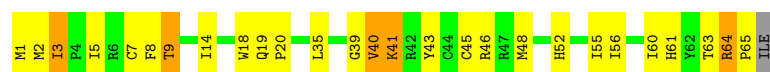
• Molecule 6: DNA-DIRECTED RNA POLYMERASE SUBUNIT N

Chain AO: 

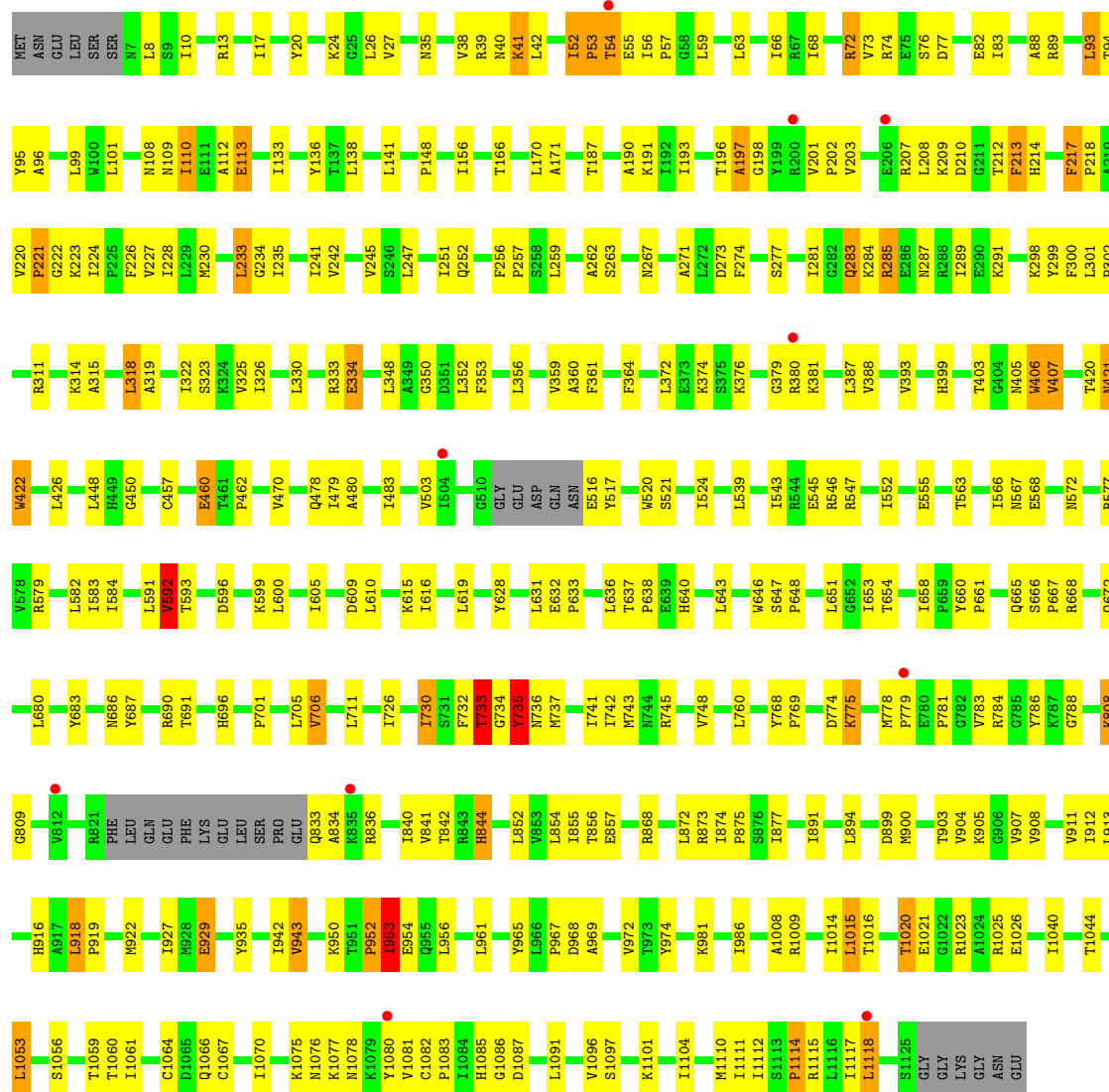


• Molecule 6: DNA-DIRECTED RNA POLYMERASE SUBUNIT N

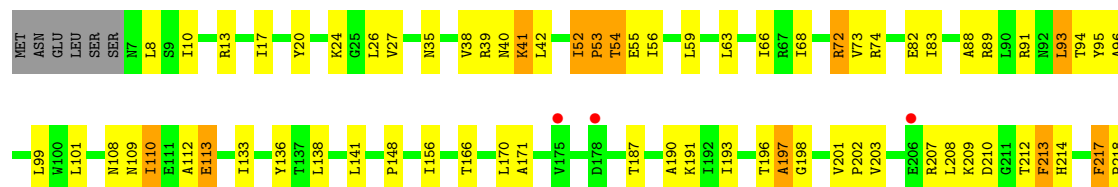
Chain BN: 

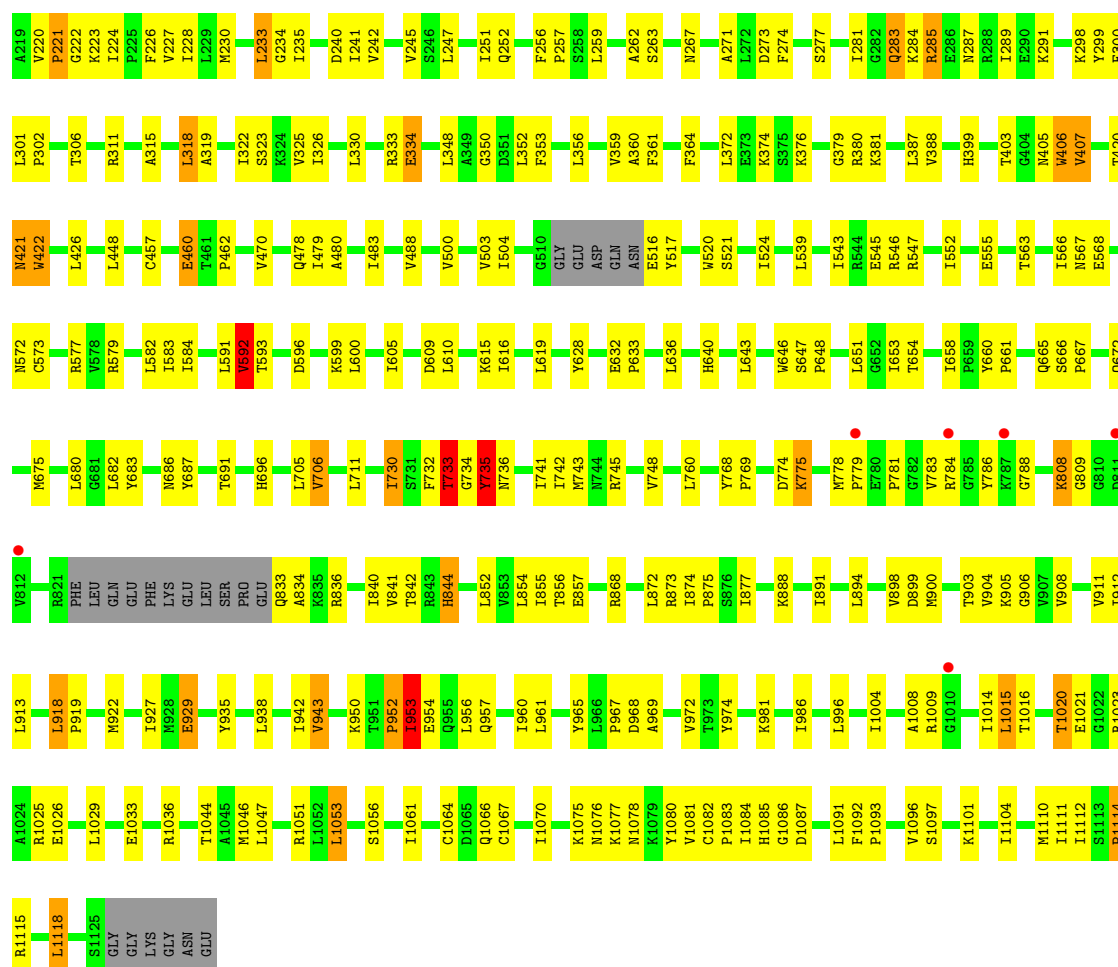


### • Molecule 7: DNA-DIRECTED RNA POLYMERASE

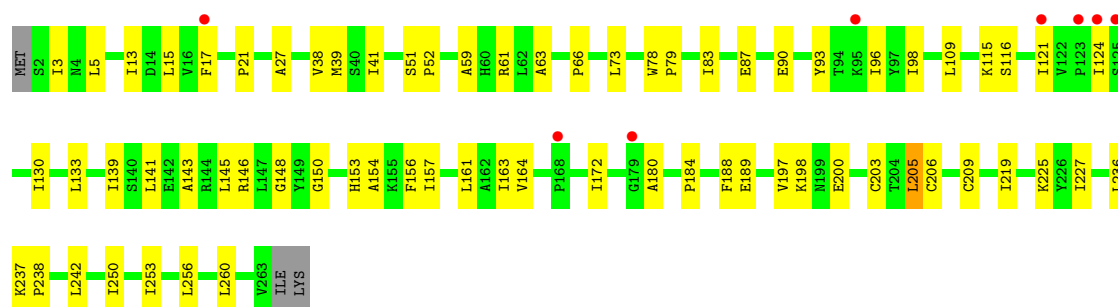
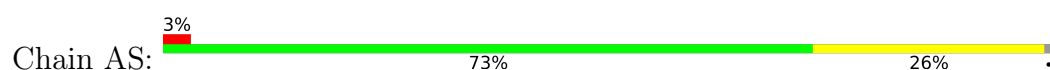


### • Molecule 7: DNA-DIRECTED RNA POLYMERASE

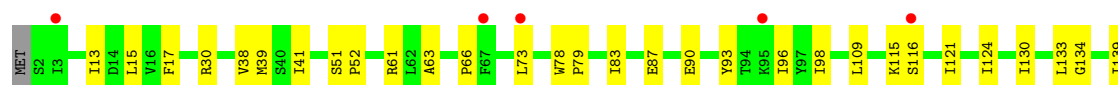
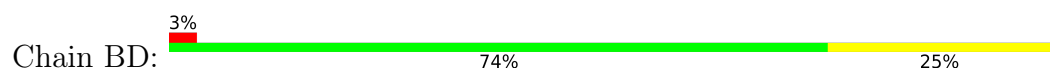


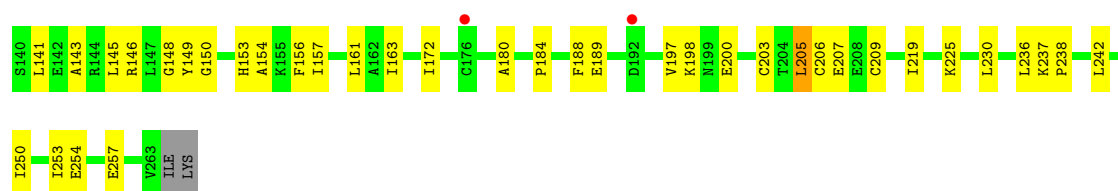


### • Molecule 8: DNA-DIRECTED RNA POLYMERASE SUBUNIT D

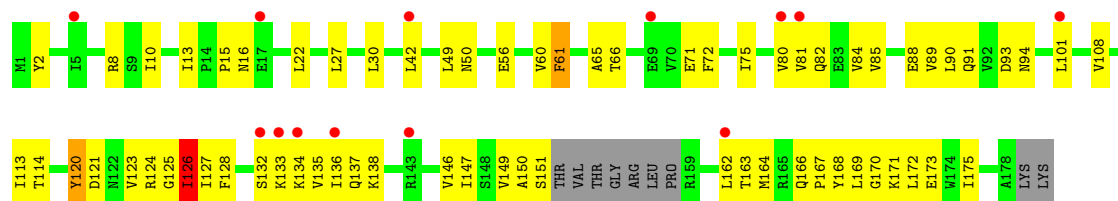


### • Molecule 8: DNA-DIRECTED RNA POLYMERASE SUBUNIT D

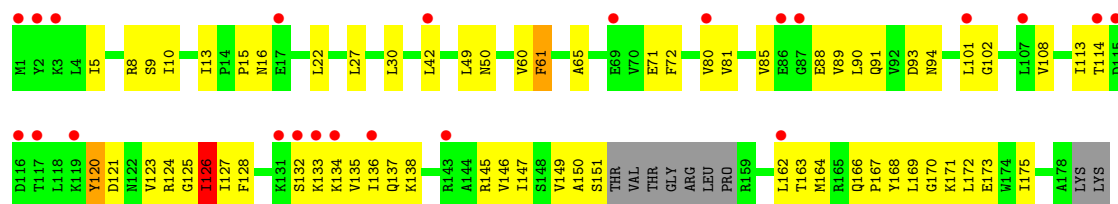




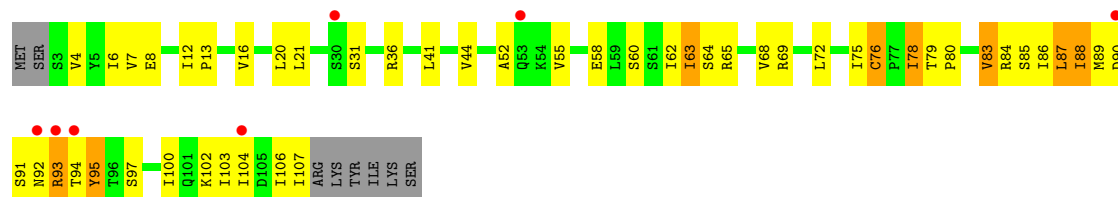
• Molecule 9: RNA POLYMERASE SUBUNIT 7



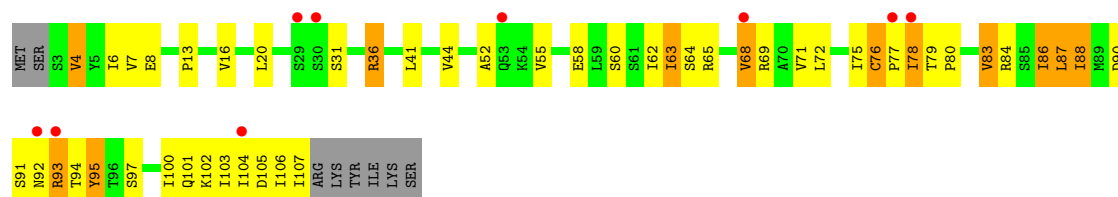
• Molecule 9: RNA POLYMERASE SUBUNIT 7



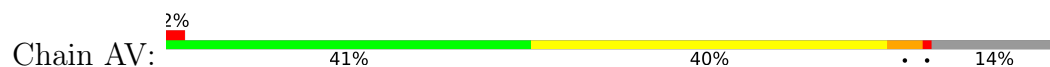
• Molecule 10: RNA POLYMERASE SUBUNIT 4

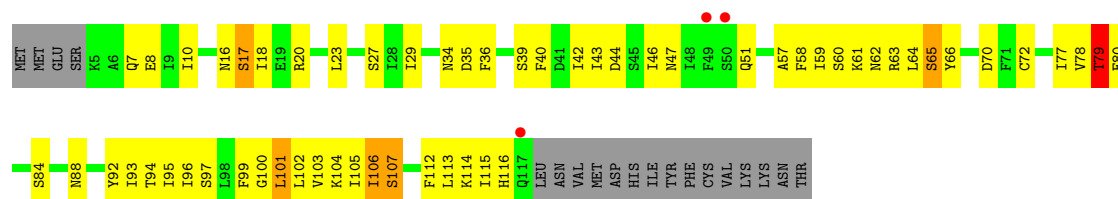


• Molecule 10: RNA POLYMERASE SUBUNIT 4

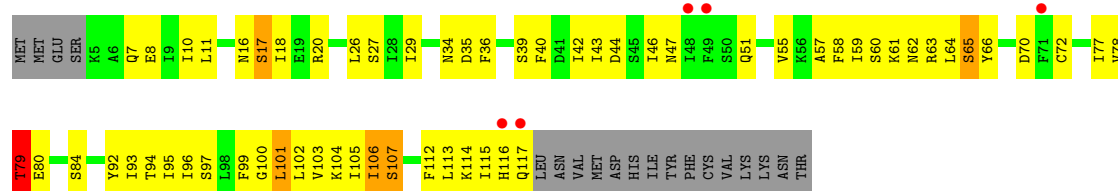
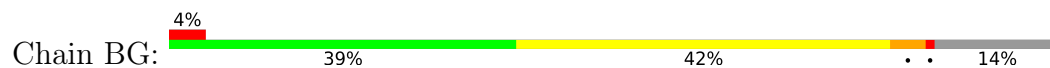


• Molecule 11: RNA POLYMERASE SUBUNIT 8

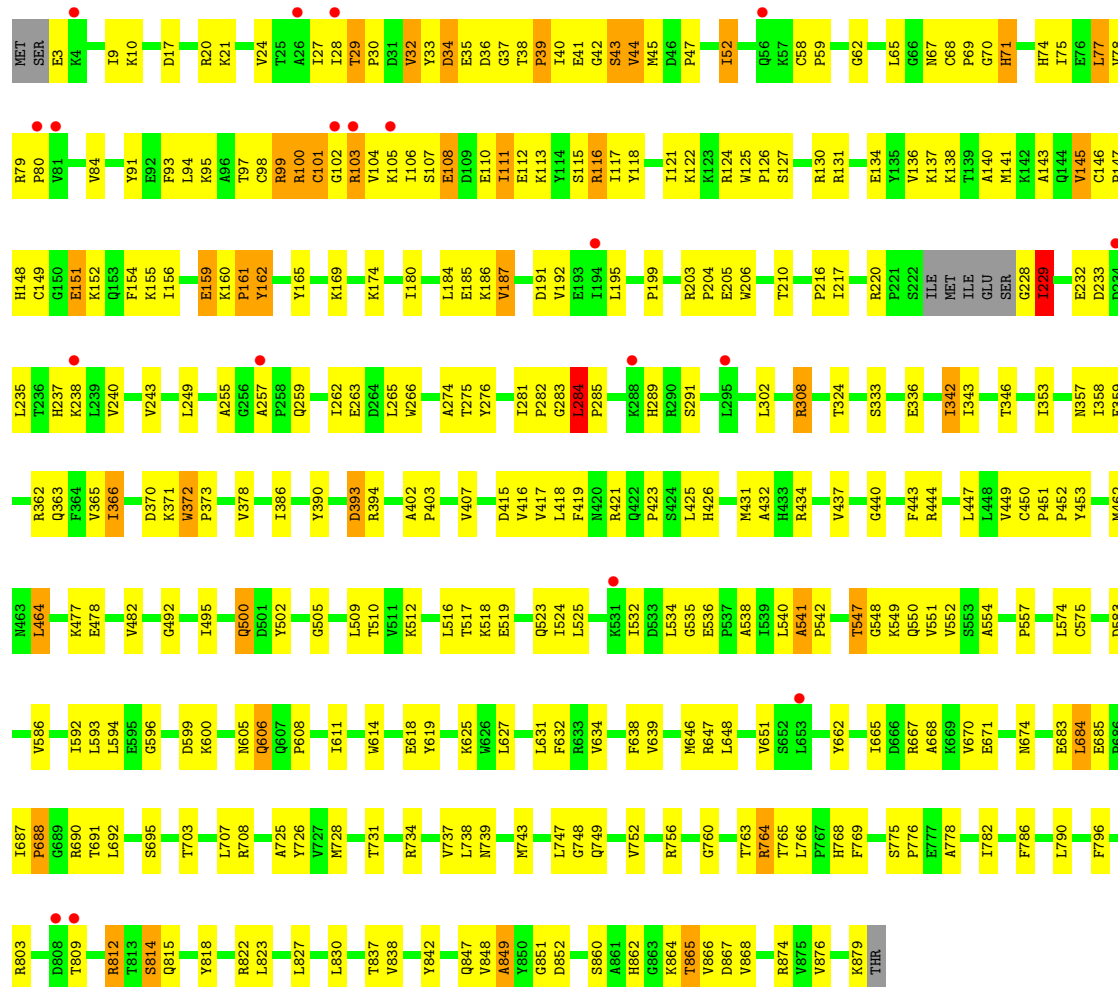




• Molecule 11: RNA POLYMERASE SUBUNIT 8



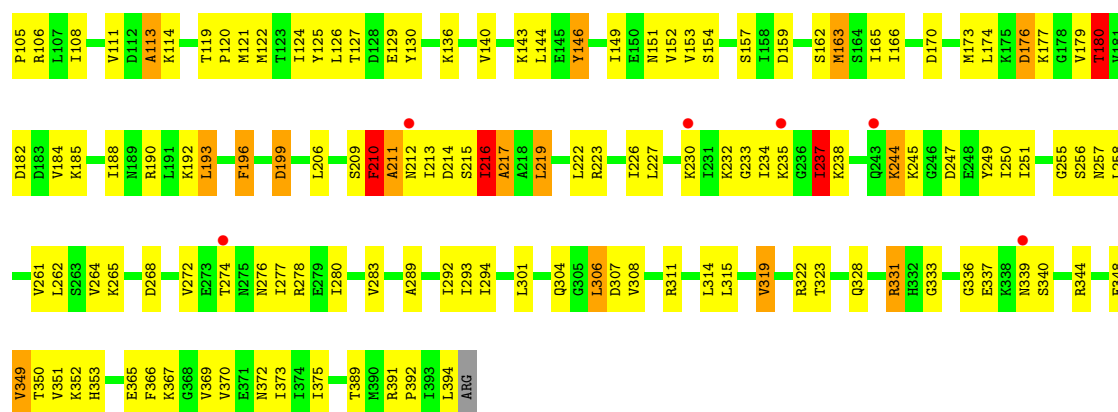
• Molecule 12: DNA-DIRECTED RNA POLYMERASE



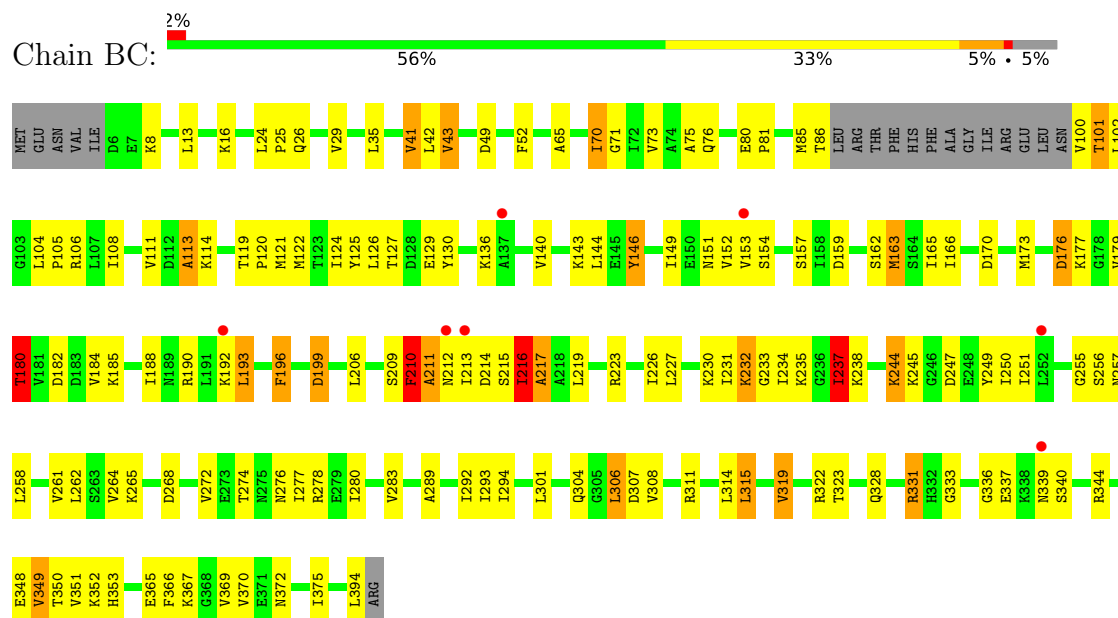
• Molecule 12: DNA-DIRECTED RNA POLYMERASE



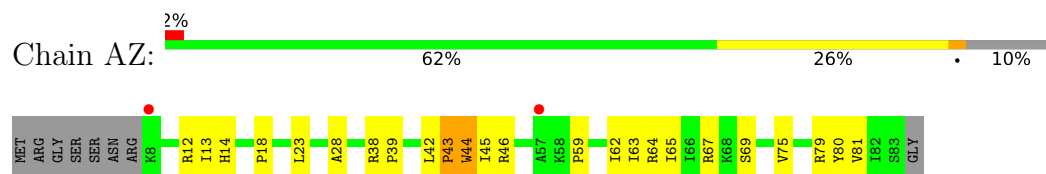




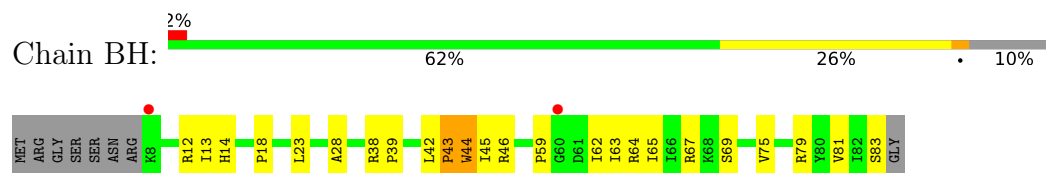
• Molecule 14: DNA-DIRECTED RNA POLYMERASE SUBUNIT A”



• Molecule 15: DNA-DIRECTED RNA POLYMERASE SUBUNIT H



• Molecule 15: DNA-DIRECTED RNA POLYMERASE SUBUNIT H



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	134.13Å 199.42Å 214.25Å 90.00° 103.54° 90.00°	Depositor
Resolution (Å)	50.42 – 4.32 50.42 – 4.32	Depositor EDS
% Data completeness (in resolution range)	88.6 (50.42-4.32) 88.6 (50.42-4.32)	Depositor EDS
$R_{merge}$	0.21	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.42 (at 4.29Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, $R_{free}$	0.292 , 0.310 0.292 , 0.303	Depositor DCC
$R_{free}$ test set	3309 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	105.4	Xtriage
Anisotropy	0.633	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 125.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.44$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.81	EDS
Total number of atoms	111598	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	171.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.90% 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: MG, ZN, SF4

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	AC	0.42	0/288	1.18	1/441 (0.2%)
1	BR	0.43	0/309	1.16	1/473 (0.2%)
2	AD	0.47	0/355	1.18	6/547 (1.1%)
2	BS	0.46	0/379	1.14	5/584 (0.9%)
3	AI	0.21	0/682	0.40	0/921
3	BK	0.22	0/682	0.41	0/921
4	AJ	0.28	0/423	0.47	0/566
4	BQ	0.29	0/432	0.44	0/578
5	AM	0.21	0/717	0.37	0/968
5	BL	0.23	0/717	0.38	0/968
6	AO	0.22	0/532	0.42	0/718
6	BN	0.23	0/532	0.43	0/718
7	AR	0.21	0/8923	0.40	0/12071
7	BB	0.21	0/8923	0.40	0/12071
8	AS	0.21	0/2123	0.37	0/2870
8	BD	0.21	0/2123	0.37	0/2870
9	AT	0.21	0/1379	0.39	0/1861
9	BE	0.21	0/1379	0.39	0/1861
10	AU	0.21	0/836	0.41	0/1133
10	BF	0.21	0/836	0.42	0/1133
11	AV	0.22	0/913	0.40	0/1224
11	BG	0.23	0/913	0.41	0/1224
12	AW	0.23	0/7108	0.42	0/9618
12	BA	0.23	0/7108	0.42	0/9618
13	AX	0.23	0/365	0.42	0/489
13	BP	0.23	0/365	0.42	0/489
14	AY	0.22	0/2930	0.44	0/3944
14	BC	0.21	0/2930	0.44	0/3944
15	AZ	0.20	0/638	0.38	0/864
15	BH	0.21	0/638	0.39	0/864
All	All	0.23	0/56478	0.45	13/76551 (0.0%)

There are no bond length outliers.

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	AD	13	DG	O4'-C1'-N9	7.03	112.92	108.00
2	BS	13	DG	O4'-C1'-N9	6.96	112.87	108.00
2	BS	15	DT	O4'-C1'-N1	6.27	112.39	108.00
2	AD	15	DT	O4'-C1'-N1	6.18	112.32	108.00
2	BS	13	DG	C1'-O4'-C4'	-5.78	104.32	110.10

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	AC	260	151	151	4	0
1	BR	279	162	162	6	0
2	AD	315	170	170	36	0
2	BS	336	181	181	25	0
3	AI	673	717	716	16	0
3	BK	673	717	716	12	0
4	AJ	417	413	413	38	0
4	BQ	426	419	419	37	0
5	AM	707	742	739	13	0
5	BL	707	742	739	12	0
6	AO	521	537	535	29	0
6	BN	521	537	535	26	0
7	AR	8756	8909	8888	285	0
7	BB	8756	8909	8888	290	0
8	AS	2087	2128	2125	53	0
8	BD	2087	2128	2125	49	0
9	AT	1359	1413	1411	60	0
9	BE	1359	1412	1411	53	0
10	AU	827	840	839	53	0
10	BF	827	840	839	51	0
11	AV	901	915	912	69	0
11	BG	901	915	912	68	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	AW	6957	7030	7013	324	0
12	BA	6957	7030	7013	316	0
13	AX	357	387	387	18	0
13	BP	357	387	387	15	0
14	AY	2906	3068	3063	168	0
14	BC	2906	3068	3063	149	0
15	AZ	624	660	658	24	0
15	BH	624	660	658	25	0
16	AO	1	0	0	0	0
16	AR	1	0	0	0	0
16	AW	3	0	0	0	0
16	AX	1	0	0	0	0
16	BA	3	0	0	0	0
16	BB	1	0	0	0	0
16	BN	1	0	0	0	0
16	BP	1	0	0	0	0
17	AS	7	0	0	2	0
17	BD	7	0	0	2	0
18	AW	1	0	0	0	0
18	BA	1	0	0	0	0
All	All	55411	56187	56068	2178	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 20.

The worst 5 of 2178 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)
10:AU:93:ARG:CB	10:AU:94:THR:HA	1.85	1.06
2:AD:14:DA:OP1	14:AY:348:GLU:HG2	1.54	1.05
10:BF:93:ARG:CB	10:BF:94:THR:HA	1.85	1.05
7:AR:221:PRO:HB2	7:AR:222:GLY:HA2	1.37	1.05
7:BB:221:PRO:HB2	7:BB:222:GLY:HA2	1.38	1.01

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	
3	AI	82/95 (86%)	68 (83%)	9 (11%)	5 (6%)	1	21
3	BK	82/95 (86%)	68 (83%)	9 (11%)	5 (6%)	1	21
4	AJ	47/104 (45%)	39 (83%)	4 (8%)	4 (8%)	1	14
4	BQ	48/104 (46%)	39 (81%)	4 (8%)	5 (10%)	0	10
5	AM	89/92 (97%)	82 (92%)	5 (6%)	2 (2%)	7	43
5	BL	89/92 (97%)	81 (91%)	6 (7%)	2 (2%)	7	43
6	AO	63/66 (96%)	47 (75%)	9 (14%)	7 (11%)	0	9
6	BN	63/66 (96%)	47 (75%)	9 (14%)	7 (11%)	0	9
7	AR	1097/1131 (97%)	925 (84%)	129 (12%)	43 (4%)	3	30
7	BB	1097/1131 (97%)	925 (84%)	130 (12%)	42 (4%)	3	31
8	AS	260/265 (98%)	218 (84%)	36 (14%)	6 (2%)	7	42
8	BD	260/265 (98%)	218 (84%)	36 (14%)	6 (2%)	7	42
9	AT	167/180 (93%)	150 (90%)	15 (9%)	2 (1%)	14	55
9	BE	167/180 (93%)	150 (90%)	15 (9%)	2 (1%)	14	55
10	AU	103/113 (91%)	79 (77%)	18 (18%)	6 (6%)	2	22
10	BF	103/113 (91%)	79 (77%)	18 (18%)	6 (6%)	2	22
11	AV	111/132 (84%)	80 (72%)	25 (22%)	6 (5%)	2	24
11	BG	111/132 (84%)	80 (72%)	25 (22%)	6 (5%)	2	24
12	AW	868/880 (99%)	716 (82%)	121 (14%)	31 (4%)	4	33
12	BA	868/880 (99%)	717 (83%)	121 (14%)	30 (4%)	4	33
13	AX	42/48 (88%)	29 (69%)	10 (24%)	3 (7%)	1	18
13	BP	42/48 (88%)	29 (69%)	10 (24%)	3 (7%)	1	18
14	AY	372/395 (94%)	303 (82%)	50 (13%)	19 (5%)	2	25
14	BC	372/395 (94%)	303 (82%)	50 (13%)	19 (5%)	2	25
15	AZ	74/84 (88%)	64 (86%)	6 (8%)	4 (5%)	2	24

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
15	BH	74/84 (88%)	64 (86%)	6 (8%)	4 (5%)	2	24
All	All	6751/7170 (94%)	5600 (83%)	876 (13%)	275 (4%)	3	29

5 of 275 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	AI	55	ASN
4	AJ	37	SER
4	AJ	58	LYS
4	AJ	59	ILE
6	AO	3	ILE

### 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	
3	AI	74/83 (89%)	74 (100%)	0	100	100
3	BK	74/83 (89%)	74 (100%)	0	100	100
4	AJ	47/96 (49%)	43 (92%)	4 (8%)	12	42
4	BQ	48/96 (50%)	44 (92%)	4 (8%)	12	42
5	AM	79/80 (99%)	79 (100%)	0	100	100
5	BL	79/80 (99%)	79 (100%)	0	100	100
6	AO	59/60 (98%)	57 (97%)	2 (3%)	40	68
6	BN	59/60 (98%)	57 (97%)	2 (3%)	40	68
7	AR	951/975 (98%)	915 (96%)	36 (4%)	36	65
7	BB	951/975 (98%)	915 (96%)	36 (4%)	36	65
8	AS	235/238 (99%)	234 (100%)	1 (0%)	92	95
8	BD	235/238 (99%)	234 (100%)	1 (0%)	92	95
9	AT	150/158 (95%)	143 (95%)	7 (5%)	29	60
9	BE	150/158 (95%)	143 (95%)	7 (5%)	29	60
10	AU	99/107 (92%)	92 (93%)	7 (7%)	16	48

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
10	BF	99/107 (92%)	92 (93%)	7 (7%)	16	48
11	AV	106/125 (85%)	102 (96%)	4 (4%)	36	65
11	BG	106/125 (85%)	102 (96%)	4 (4%)	36	65
12	AW	758/766 (99%)	717 (95%)	41 (5%)	24	57
12	BA	758/766 (99%)	717 (95%)	41 (5%)	24	57
13	AX	40/43 (93%)	40 (100%)	0	100	100
13	BP	40/43 (93%)	40 (100%)	0	100	100
14	AY	324/341 (95%)	306 (94%)	18 (6%)	23	56
14	BC	324/341 (95%)	306 (94%)	18 (6%)	23	56
15	AZ	69/75 (92%)	68 (99%)	1 (1%)	69	85
15	BH	69/75 (92%)	68 (99%)	1 (1%)	69	85
All	All	5983/6294 (95%)	5741 (96%)	242 (4%)	34	63

5 of 242 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
14	AY	210	PHE
12	BA	162	TYR
9	BE	126	ILE
14	AY	232	LYS
12	BA	32	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
12	AW	357	ASN
12	AW	485	ASN
12	BA	500	GLN
12	AW	237	HIS
12	BA	485	ASN

### 5.3.3 RNA ⓘ

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 16 ligands modelled in this entry, 14 are 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	$\# Z  > 2$
17	SF4	AS	1001	8	0,9,12	0.00	-	-		
17	SF4	BD	1001	8	0,9,12	0.00	-	-		

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 ⓘ

### 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	AC	13/14 (92%)	1.16	3 (23%) 0 0	263, 280, 311, 312	0
1	BR	14/14 (100%)	0.91	2 (14%) 2 2	228, 239, 287, 301	0
2	AD	15/16 (93%)	1.00	2 (13%) 3 3	264, 278, 302, 313	0
2	BS	16/16 (100%)	0.62	2 (12%) 4 4	216, 237, 276, 283	0
3	AI	84/95 (88%)	-0.05	0 100 100	139, 154, 175, 183	0
3	BK	84/95 (88%)	-0.08	0 100 100	122, 150, 173, 184	0
4	AJ	49/104 (47%)	0.72	3 (6%) 21 15	219, 247, 277, 291	0
4	BQ	50/104 (48%)	0.62	3 (6%) 22 16	216, 237, 255, 263	0
5	AM	91/92 (98%)	0.03	3 (3%) 46 35	159, 184, 196, 200	0
5	BL	91/92 (98%)	0.04	2 (2%) 62 51	139, 164, 177, 179	0
6	AO	65/66 (98%)	-0.03	0 100 100	178, 193, 215, 220	0
6	BN	65/66 (98%)	-0.06	0 100 100	137, 167, 202, 206	0
7	AR	1103/1131 (97%)	0.04	10 (0%) 84 76	132, 158, 195, 219	0
7	BB	1103/1131 (97%)	-0.00	9 (0%) 86 79	117, 146, 186, 210	0
8	AS	262/265 (98%)	0.24	8 (3%) 49 37	166, 202, 222, 228	0
8	BD	262/265 (98%)	0.20	7 (2%) 54 43	138, 170, 213, 245	0
9	AT	171/180 (95%)	0.53	13 (7%) 14 10	161, 232, 291, 312	0
9	BE	171/180 (95%)	0.78	23 (13%) 3 3	160, 222, 294, 312	0
10	AU	105/113 (92%)	0.31	7 (6%) 18 12	194, 285, 328, 341	0
10	BF	105/113 (92%)	0.39	9 (8%) 10 8	182, 261, 303, 318	0
11	AV	113/132 (85%)	0.20	3 (2%) 54 43	164, 199, 234, 246	0
11	BG	113/132 (85%)	0.31	5 (4%) 34 26	140, 182, 218, 230	0
12	AW	872/880 (99%)	0.12	19 (2%) 62 51	131, 162, 230, 265	0
12	BA	872/880 (99%)	0.02	13 (1%) 73 63	116, 145, 198, 227	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
13	AX	44/48 (91%)	-0.04	0 100 100	164, 194, 206, 210	0
13	BP	44/48 (91%)	-0.14	0 100 100	132, 188, 206, 210	0
14	AY	376/395 (95%)	0.12	6 (1%) 72 61	148, 168, 222, 237	0
14	BC	376/395 (95%)	0.11	7 (1%) 66 56	125, 154, 209, 230	0
15	AZ	76/84 (90%)	-0.05	2 (2%) 56 44	157, 185, 200, 206	0
15	BH	76/84 (90%)	-0.21	2 (2%) 56 44	134, 156, 174, 184	0
All	All	6881/7230 (95%)	0.12	163 (2%) 59 47	116, 164, 251, 341	0

The worst 5 of 163 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
9	BE	132	SER	6.1
14	BC	212	ASN	5.6
12	BA	29	THR	5.3
15	AZ	8	LYS	5.2
10	AU	92	ASN	4.7

## 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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
16	ZN	BN	100	1/1	0.78	0.23	153,153,153,153	0
16	ZN	AX	101	1/1	0.82	0.10	204,204,204,204	0
16	ZN	AW	901	1/1	0.83	0.07	183,183,183,183	0
16	ZN	BP	101	1/1	0.85	0.09	200,200,200,200	0

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
16	ZN	BA	901	1/1	0.88	0.08	161,161,161,161	0
18	MG	BA	904	1/1	0.88	0.37	117,117,117,117	0
16	ZN	AO	100	1/1	0.92	0.25	188,188,188,188	0
18	MG	AW	904	1/1	0.92	0.29	132,132,132,132	0
16	ZN	AW	903	1/1	0.93	0.11	134,134,134,134	0
16	ZN	BA	902	1/1	0.94	0.17	209,209,209,209	0
17	SF4	BD	1001	7/8	0.96	0.23	149,150,154,157	0
16	ZN	AR	1300	1/1	0.96	0.12	193,193,193,193	0
16	ZN	AW	902	1/1	0.96	0.28	245,245,245,245	0
17	SF4	AS	1001	7/8	0.97	0.24	186,189,193,194	0
16	ZN	BA	903	1/1	0.97	0.10	118,118,118,118	0
16	ZN	BB	1300	1/1	0.97	0.10	177,177,177,177	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.