



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 29, 2023 – 02:34 PM EDT

PDB ID : 3PO3  
Title : Arrested RNA Polymerase II reactivation intermediate  
Authors : Cheung, A.C.M.; Cramer, P.  
Deposited on : 2010-11-21  
Resolution : 3.30 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

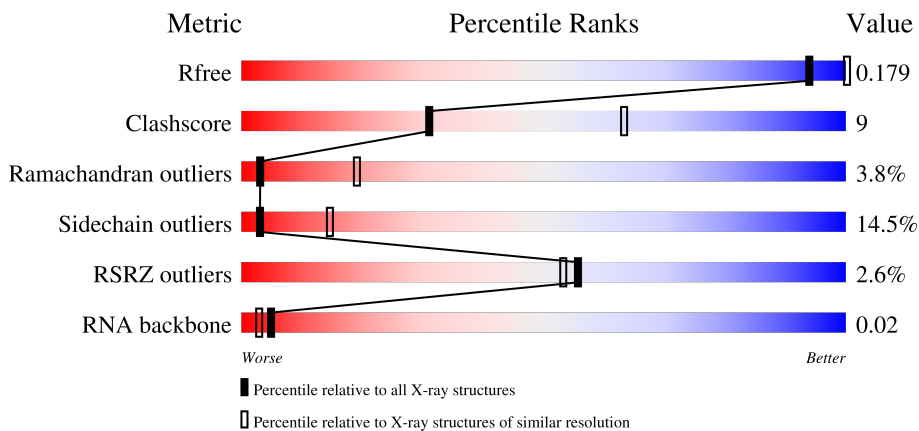
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)
RNA backbone	3102	1117 (3.70-2.90)

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 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	1733	 56% 21% 5% 18%
2	B	1224	 66% 22% 9%
3	C	318	 53% 25% 5% 16%
4	D	221	 53% 20% 6% 20%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	215	
6	F	155	
7	G	171	
8	H	146	
9	I	122	
10	J	70	
11	K	120	
12	L	70	
13	N	14	
14	P	5	
15	S	178	
16	T	27	

## 2 Entry composition [i](#)

There are 21 unique types of molecules in this entry. The entry contains 33008 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 II subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1426	11214	7069	1959	2124	62	0	0	0

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	1108	8810	5580	1541	1634	55	0	0	0

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	266	2095	1317	348	417	13	0	0	0

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	177	1417	876	252	287	2	0	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	E	214	1752	1111	309	321	11	0	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	84	679	434	115	127	3	0	0	0

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	G	171	1340	861	222	249	8	0	0	0

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	133	1068	673	180	211	4	0	0	0

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
9	I	119	971	596	179	186	10	0	0	0

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
10	J	65	532	339	93	94	6	0	0	0

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
11	K	114	919	590	156	171	2	0	0	0

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
12	L	46	364	224	72	64	4	0	0	0

- Molecule 13 is a DNA chain called DNA non-template strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
13	N	7	144	69	30	39	6	0	0	0

- Molecule 14 is a RNA chain called RNA product strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
14	P	5	100	45	15	35	5	0	0	0

- Molecule 15 is a protein called Transcription elongation factor S-II.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
15	S	164	1294	809	230	247	8	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
S	290	ALA	ASP	engineered mutation	UNP P07273
S	291	ALA	GLU	engineered mutation	UNP P07273

- Molecule 16 is a DNA chain called DNA template strand.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	Br	C	N	O	P			
16	T	13	266	1	126	44	82	13	0	0	0

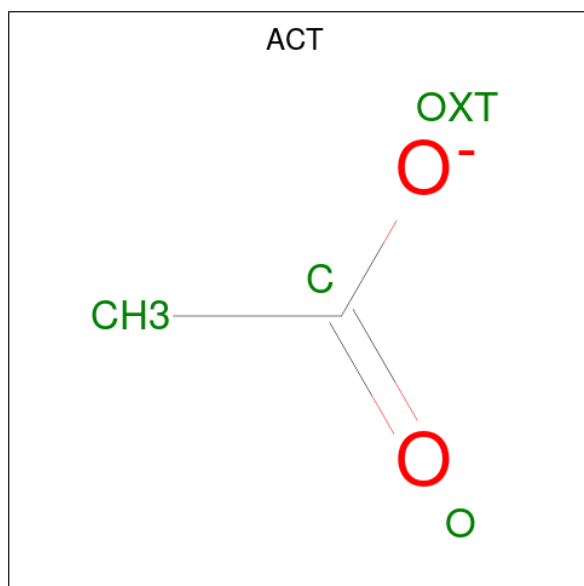
- Molecule 17 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	A	2	Total 2	Zn 2	0	0
17	B	1	Total 1	Zn 1	0	0
17	C	1	Total 1	Zn 1	0	0
17	I	2	Total 2	Zn 2	0	0
17	J	1	Total 1	Zn 1	0	0
17	L	1	Total 1	Zn 1	0	0
17	S	1	Total 1	Zn 1	0	0

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

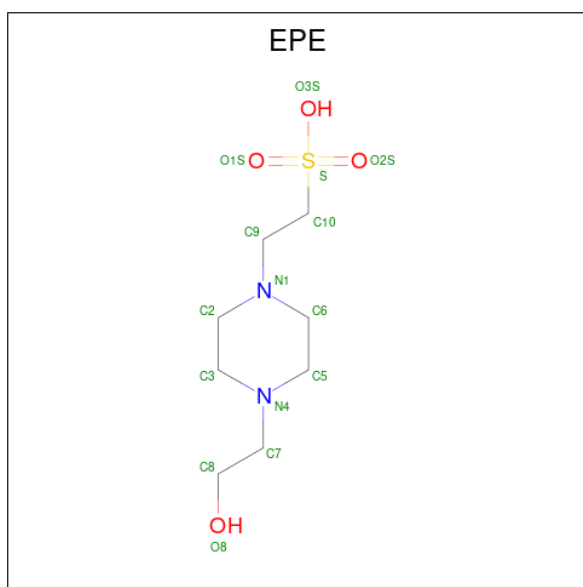
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
18	A	1	Total Mg 1 1	0	0

- Molecule 19 is ACETATE ION (three-letter code: ACT) (formula:  $C_2H_3O_2$ ).



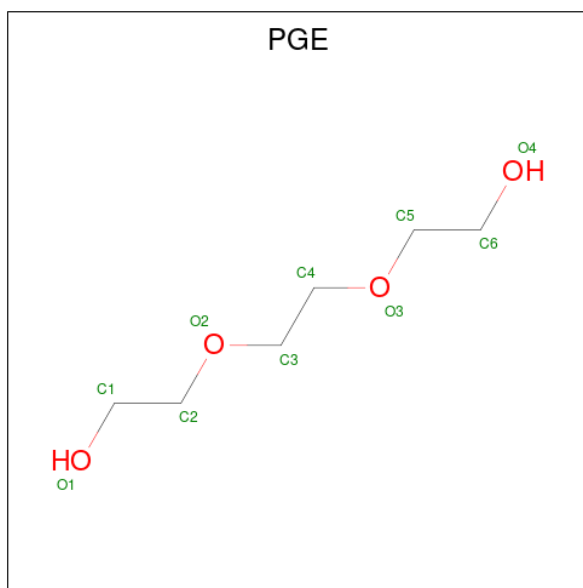
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
19	A	1	Total C O 4 2 2	0	0
19	B	1	Total C O 4 2 2	0	0

- Molecule 20 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula:  $C_8H_{18}N_2O_4S$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
20	A	1	15	8	2	4	1	0	0

- Molecule 21 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).



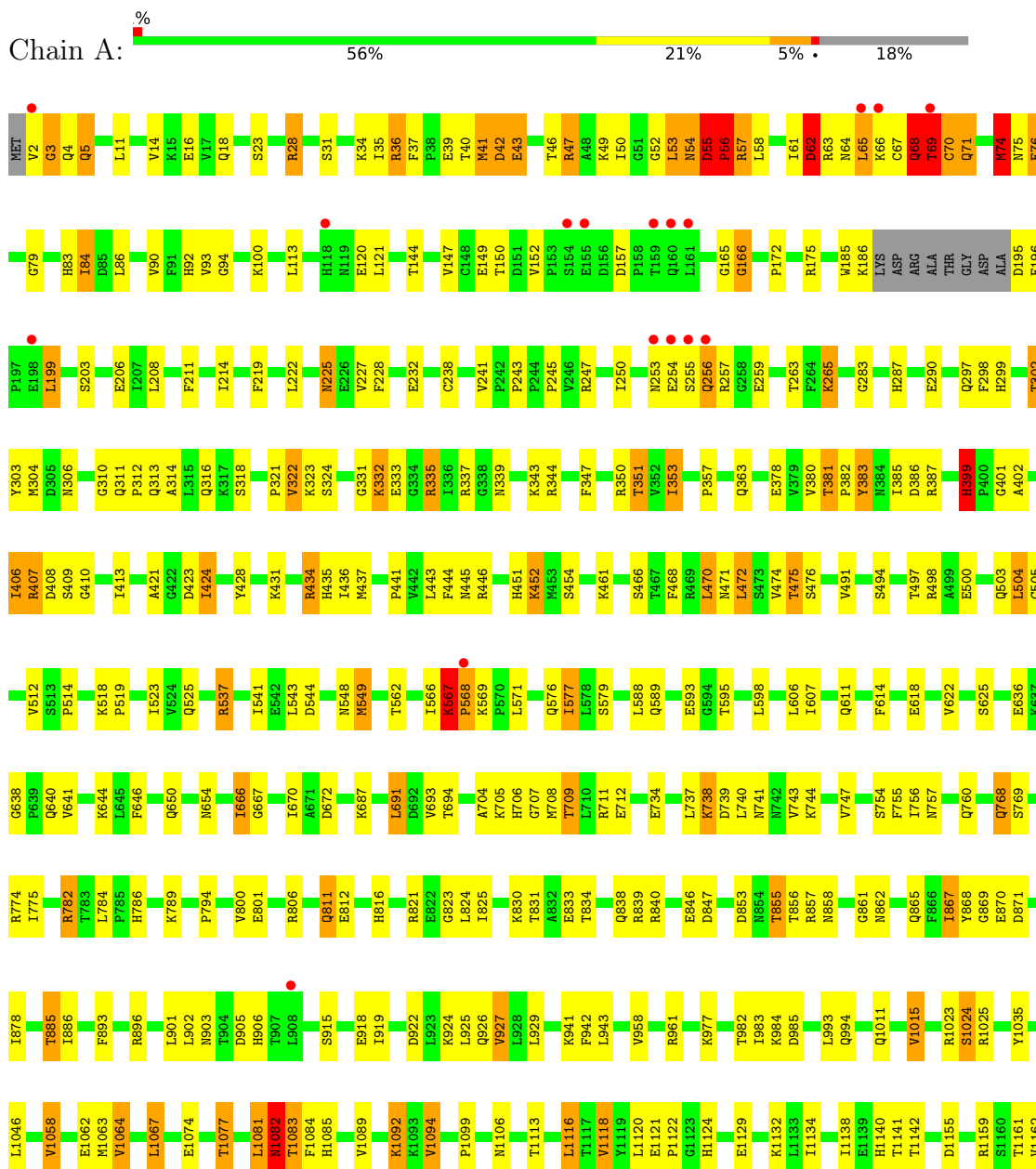
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
21	L	1	10	6	4	0	0

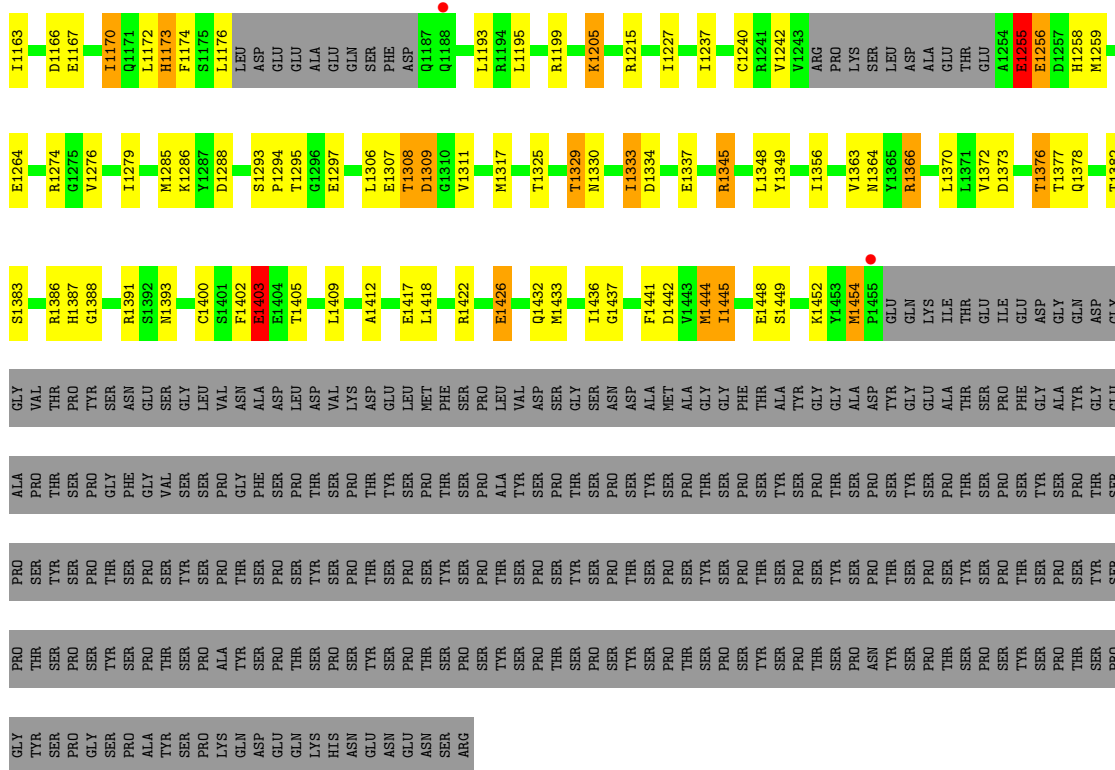


### 3 Residue-property plots [i](#)

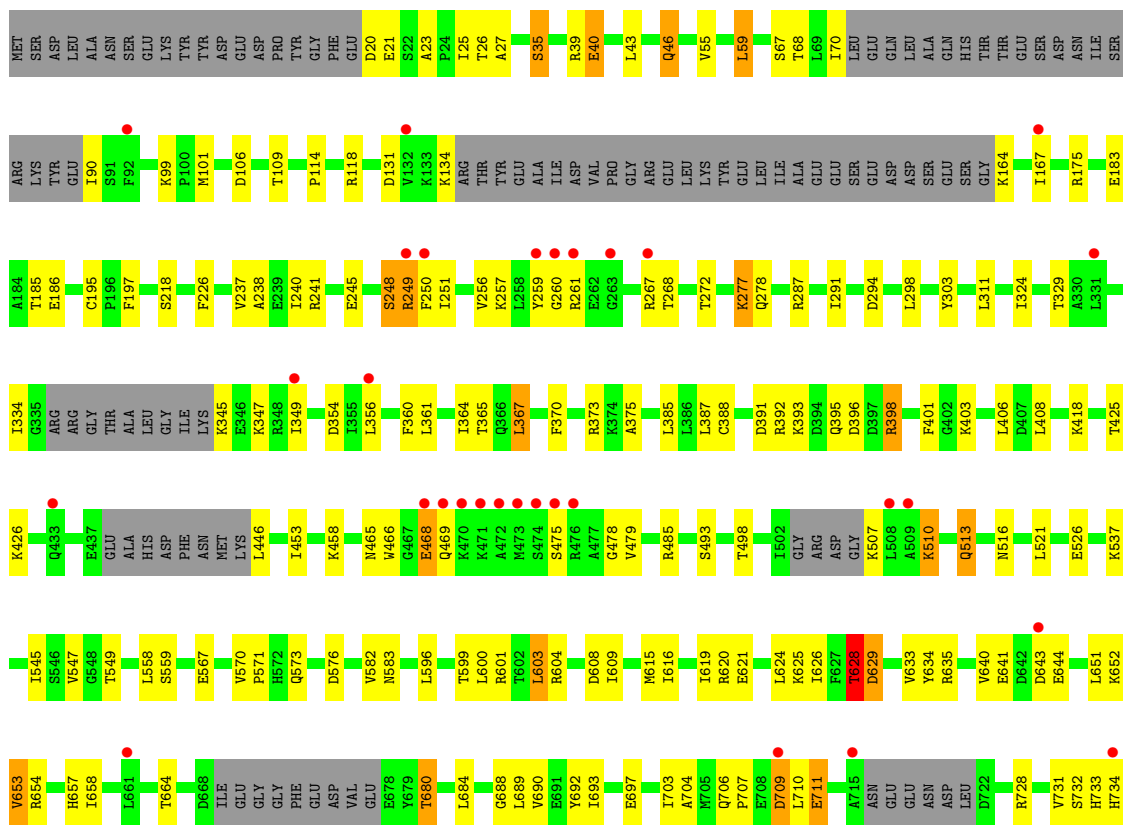
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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.

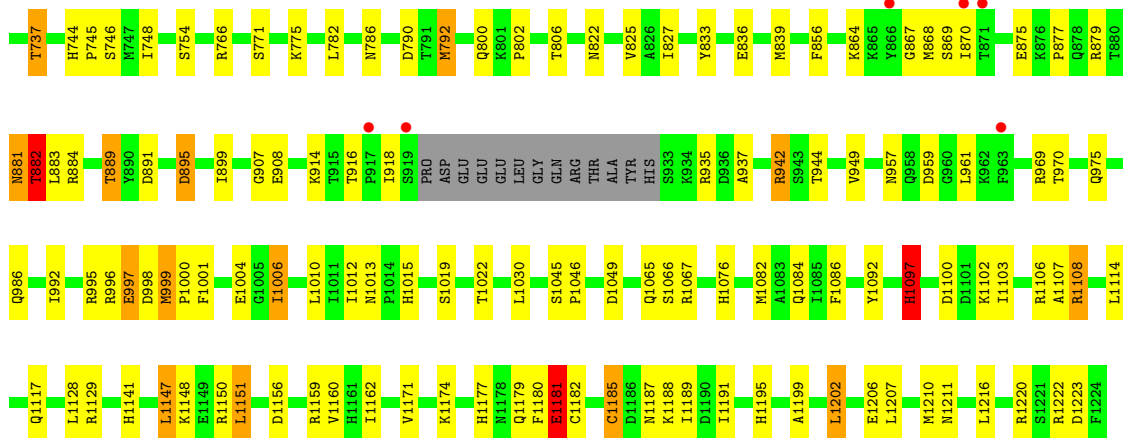
- Molecule 1: DNA-directed RNA polymerase II subunit RPB1



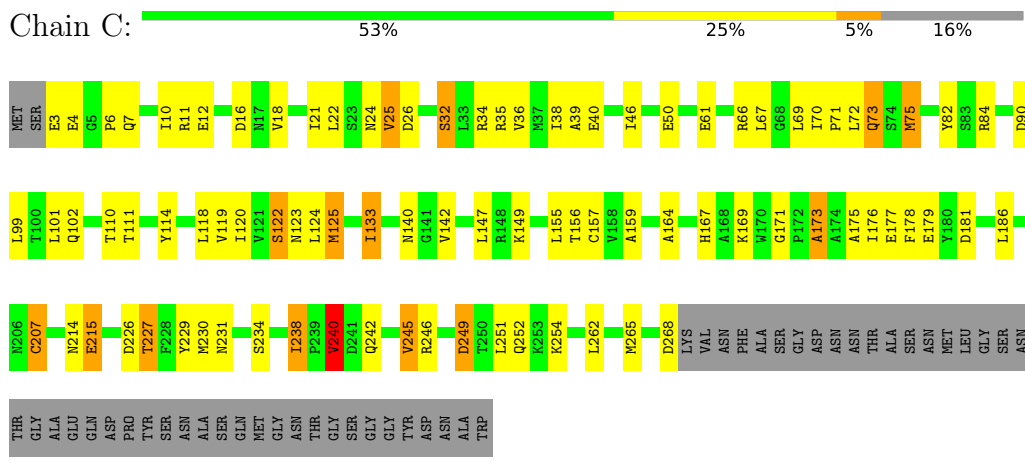


● Molecule 2: DNA-directed RNA polymerase II subunit RPB2

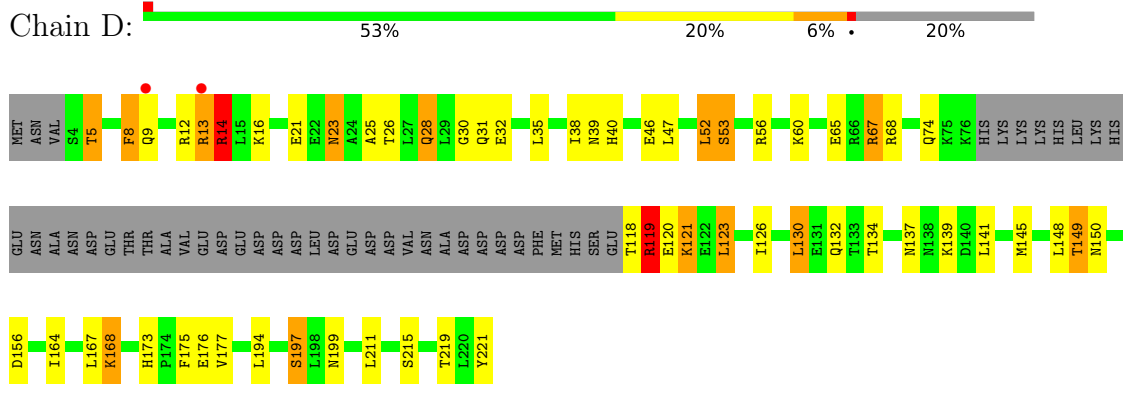




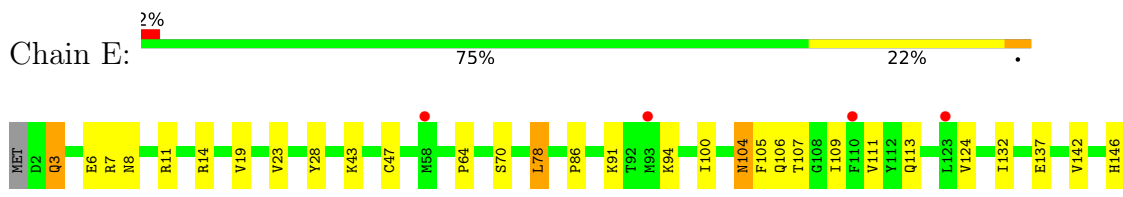
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3



• Molecule 4: DNA-directed RNA polymerase II subunit RPB4



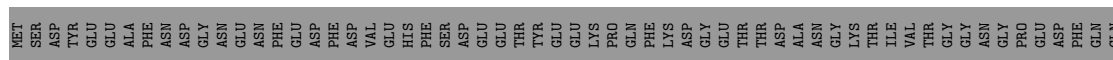
• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1





- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F: 37% 15% 46%



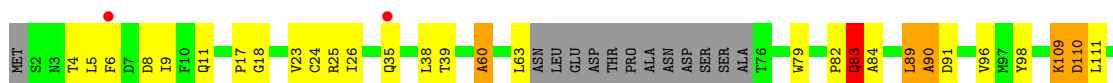
- Molecule 7: DNA-directed RNA polymerase II subunit RPB7

Chain G: % 65% 26% 9%



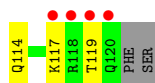
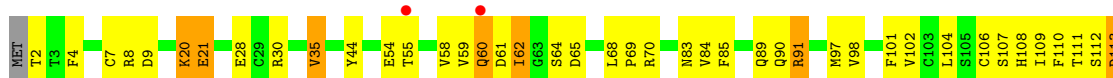
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H: 3% 64% 21% 5% 9%



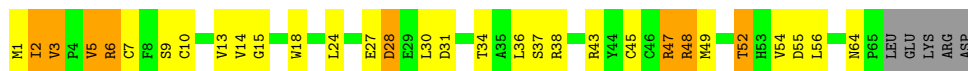
- Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain I: 5% 61% 31% 6%

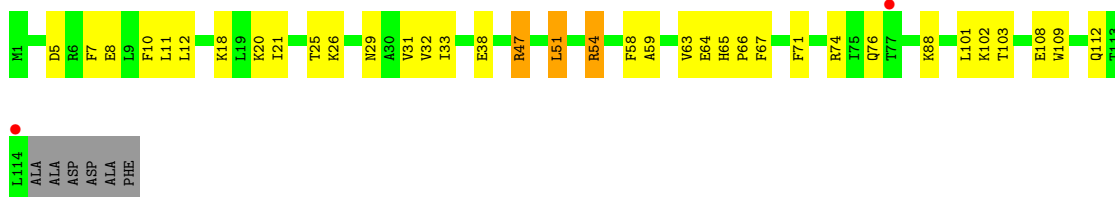


- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

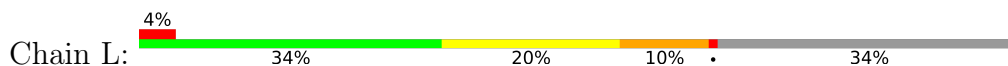
Chain J: 49% 33% 11% 7%



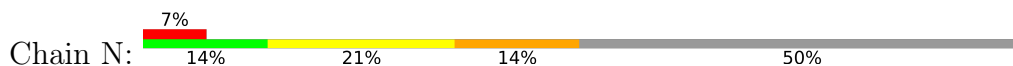
- Molecule 11: DNA-directed RNA polymerase II subunit RPB1



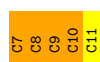
- Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4



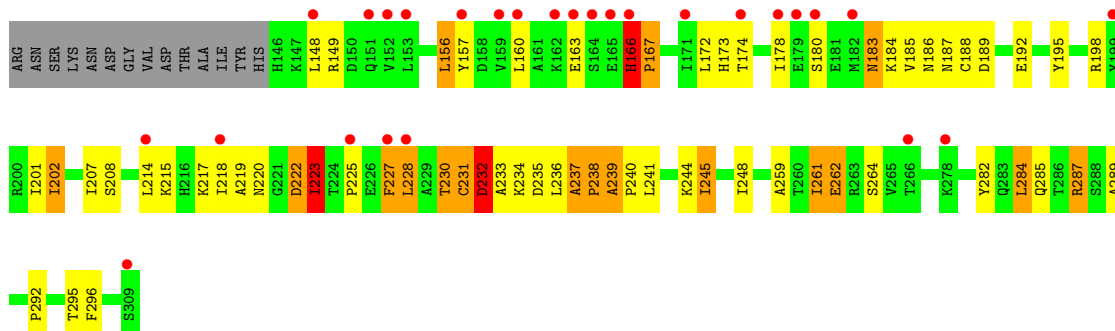
- Molecule 13: DNA non-template strand




- Molecule 14: RNA product strand

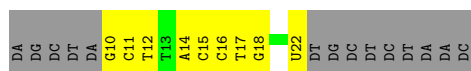


- Molecule 15: Transcription elongation factor S-II



- Molecule 16: DNA template strand

Chain T:  15% 33% 52%



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	220.02Å 395.07Å 280.43Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.97 – 3.30 49.97 – 3.30	Depositor EDS
% Data completeness (in resolution range)	(Not available) (49.97-3.30) 99.7 (49.97-3.30)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.00 (at 3.33Å)	Xtrriage
Refinement program	BUSTER 2.9.2	Depositor
R, $R_{free}$	0.161 , 0.189 0.181 , 0.179	Depositor DCC
$R_{free}$ test set	3598 reflections (1.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	111.5	Xtrriage
Anisotropy	0.234	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 116.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.018 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.026 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	33008	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	127.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.93% 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, PGE, BRU, EPE, ACT

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	A	0.54	2/11417 (0.0%)	0.85	7/15442 (0.0%)
2	B	0.51	0/8981	0.79	4/12108 (0.0%)
3	C	0.50	0/2133	0.77	1/2891 (0.0%)
4	D	0.52	0/1427	0.86	2/1914 (0.1%)
5	E	0.44	0/1788	0.71	0/2406
6	F	0.56	0/691	0.77	0/933
7	G	0.55	0/1368	0.82	0/1844
8	H	0.47	0/1086	0.80	0/1470
9	I	0.47	0/989	0.79	0/1331
10	J	0.52	0/541	0.87	1/727 (0.1%)
11	K	0.51	0/937	0.71	0/1265
12	L	0.53	0/366	0.97	1/485 (0.2%)
13	N	0.99	0/162	1.81	6/249 (2.4%)
14	P	1.26	0/109	2.55	10/166 (6.0%)
15	S	0.50	0/1317	0.79	0/1778
16	T	1.16	0/274	1.82	9/421 (2.1%)
All	All	0.53	2/33586 (0.0%)	0.85	41/45430 (0.1%)

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
1	A	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	867	ILE	CG1-CD1	5.94	1.91	1.50
1	A	56	PRO	C-N	5.29	1.46	1.34



The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
14	P	8	C	P-O3'-C3'	13.71	136.16	119.70
13	N	6	DA	O4'-C1'-N9	9.07	114.35	108.00
16	T	15	DC	O4'-C1'-N1	8.13	113.69	108.00
14	P	9	C	P-O3'-C3'	7.96	129.25	119.70
16	T	15	DC	O4'-C4'-C3'	-7.65	101.41	106.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	55	ASP	Mainchain

## 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	11214	0	11281	258	0
2	B	8810	0	8847	133	0
3	C	2095	0	2051	57	0
4	D	1417	0	1429	26	0
5	E	1752	0	1776	30	0
6	F	679	0	701	17	0
7	G	1340	0	1357	38	0
8	H	1068	0	1040	18	0
9	I	971	0	927	17	0
10	J	532	0	542	22	0
11	K	919	0	929	24	0
12	L	364	0	386	9	0
13	N	144	0	80	2	0
14	P	100	0	56	1	0
15	S	1294	0	1295	36	0
16	T	266	0	146	0	0
17	A	2	0	0	0	0
17	B	1	0	0	0	0
17	C	1	0	0	0	0
17	I	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
17	J	1	0	0	0	0
17	L	1	0	0	0	0
17	S	1	0	0	0	0
18	A	1	0	0	0	0
19	A	4	0	3	1	0
19	B	4	0	3	1	0
20	A	15	0	17	2	0
21	L	10	0	14	0	0
All	All	33008	0	32880	599	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 9.

The worst 5 of 599 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)
1:A:867:ILE:CG1	1:A:867:ILE:CD1	1.91	1.49
6:F:93:ILE:HD11	6:F:134:ILE:HD11	1.41	1.02
1:A:855:THR:HG21	1:A:857:ARG:HE	1.26	0.98
1:A:869:GLY:O	5:E:204:THR:HG21	1.69	0.93
7:G:1:MET:SD	7:G:2:PHE:N	2.44	0.90

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	1418/1733 (82%)	1255 (88%)	105 (7%)	58 (4%)	<b>3</b> <b>17</b>
2	B	1090/1224 (89%)	963 (88%)	92 (8%)	35 (3%)	<b>4</b> <b>22</b>
3	C	264/318 (83%)	238 (90%)	17 (6%)	9 (3%)	<b>3</b> <b>22</b>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	D	173/221 (78%)	153 (88%)	14 (8%)	6 (4%)	3	21
5	E	212/215 (99%)	202 (95%)	8 (4%)	2 (1%)	17	48
6	F	82/155 (53%)	77 (94%)	4 (5%)	1 (1%)	13	42
7	G	169/171 (99%)	154 (91%)	12 (7%)	3 (2%)	8	35
8	H	129/146 (88%)	101 (78%)	18 (14%)	10 (8%)	1	6
9	I	117/122 (96%)	98 (84%)	16 (14%)	3 (3%)	5	27
10	J	63/70 (90%)	54 (86%)	6 (10%)	3 (5%)	2	14
11	K	112/120 (93%)	107 (96%)	5 (4%)	0	100	100
12	L	44/70 (63%)	27 (61%)	8 (18%)	9 (20%)	0	0
15	S	162/178 (91%)	129 (80%)	20 (12%)	13 (8%)	1	6
All	All	4035/4743 (85%)	3558 (88%)	325 (8%)	152 (4%)	3	19

5 of 152 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	GLN
1	A	42	ASP
1	A	65	LEU
1	A	66	LYS
1	A	68	GLN

### 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	1246/1520 (82%)	1073 (86%)	173 (14%)	3	16
2	B	962/1061 (91%)	832 (86%)	130 (14%)	4	16
3	C	234/274 (85%)	199 (85%)	35 (15%)	3	13
4	D	157/200 (78%)	120 (76%)	37 (24%)	1	3
5	E	196/197 (100%)	178 (91%)	18 (9%)	9	31
6	F	74/137 (54%)	69 (93%)	5 (7%)	16	44

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
7	G	152/152 (100%)	125 (82%)	27 (18%)	2	8
8	H	117/128 (91%)	103 (88%)	14 (12%)	5	20
9	I	113/116 (97%)	90 (80%)	23 (20%)	1	4
10	J	60/65 (92%)	48 (80%)	12 (20%)	1	5
11	K	99/102 (97%)	87 (88%)	12 (12%)	5	20
12	L	40/57 (70%)	30 (75%)	10 (25%)	0	2
15	S	141/153 (92%)	117 (83%)	24 (17%)	2	9
All	All	3591/4162 (86%)	3071 (86%)	520 (14%)	3	14

5 of 520 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
10	J	3	VAL
11	K	25	THR
10	J	2	ILE
2	B	251	ILE
2	B	218	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 83 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	1195	HIS
5	E	5	ASN
3	C	7	GLN
3	C	167	HIS
7	G	96	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
14	P	5/5 (100%)	3 (60%)	2 (40%)

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
14	P	8	C
14	P	9	C

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type
14	P	10	C

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
14	P	7	C
14	P	8	C

## 5.4 Non-standard residues in protein, DNA, RNA chains

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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
16	BRU	T	22	16,14	18,21,22	0.70	0	26,30,33	2.15	9 (34%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
16	BRU	T	22	16,14	-	2/7/21/22	0/2/2/2

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
16	T	22	BRU	C5-C4-N3	4.86	118.93	113.34
16	T	22	BRU	C4-N3-C2	-4.70	121.27	127.35
16	T	22	BRU	N3-C2-N1	4.60	120.99	114.89
16	T	22	BRU	O4-C4-C5	-3.55	121.39	125.84
16	T	22	BRU	O4'-C1'-N1	3.16	113.51	107.86

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
16	T	22	BRU	C3'-C4'-C5'-O5'
16	T	22	BRU	O4'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 14 ligands modelled in this entry, 10 are monoatomic - leaving 4 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
19	ACT	A	1734	-	3,3,3	0.69	0	3,3,3	1.19	0
21	PGE	L	71	-	9,9,9	0.48	0	8,8,8	0.32	0
19	ACT	B	1225	-	3,3,3	0.89	0	3,3,3	1.80	1 (33%)
20	EPE	A	1735	-	15,15,15	0.60	0	18,20,20	2.12	7 (38%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
21	PGE	L	71	-	-	3/7/7/7	-
20	EPE	A	1735	-	-	3/9/19/19	0/1/1/1

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
20	A	1735	EPE	C5-N4-C3	4.76	119.54	108.83
20	A	1735	EPE	C7-N4-C3	3.68	120.64	111.23
20	A	1735	EPE	C7-N4-C5	3.30	119.68	111.23
20	A	1735	EPE	C9-N1-C2	-2.83	103.99	111.23
19	B	1225	ACT	OXT-C-CH3	2.35	124.90	115.18

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
21	L	71	PGE	O2-C3-C4-O3
21	L	71	PGE	O3-C5-C6-O4
21	L	71	PGE	O1-C1-C2-O2
20	A	1735	EPE	C10-C9-N1-C2
20	A	1735	EPE	C10-C9-N1-C6

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
19	A	1734	ACT	1	0
19	B	1225	ACT	1	0
20	A	1735	EPE	2	0

## 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	1426/1733 (82%)	-0.02	19 (1%) 77 77	67, 109, 163, 215	0
2	B	1108/1224 (90%)	0.16	36 (3%) 47 46	69, 126, 191, 217	0
3	C	266/318 (83%)	-0.11	0 100 100	86, 116, 160, 193	0
4	D	177/221 (80%)	-0.06	2 (1%) 80 81	81, 118, 183, 202	0
5	E	214/215 (99%)	-0.01	4 (1%) 66 65	88, 141, 188, 203	0
6	F	84/155 (54%)	-0.26	0 100 100	70, 90, 122, 136	0
7	G	171/171 (100%)	0.06	1 (0%) 89 90	82, 105, 148, 171	0
8	H	133/146 (91%)	0.39	4 (3%) 50 49	121, 158, 197, 205	0
9	I	119/122 (97%)	0.25	6 (5%) 28 27	123, 156, 197, 216	0
10	J	65/70 (92%)	-0.06	0 100 100	95, 117, 159, 177	0
11	K	114/120 (95%)	-0.04	2 (1%) 68 67	85, 112, 157, 173	0
12	L	46/70 (65%)	0.07	3 (6%) 18 18	93, 158, 183, 202	0
13	N	7/14 (50%)	1.26	1 (14%) 2 2	217, 222, 231, 234	0
14	P	5/5 (100%)	0.22	0 100 100	238, 239, 244, 247	0
15	S	164/178 (92%)	0.65	27 (16%) 1 2	116, 181, 240, 254	0
16	T	12/27 (44%)	0.57	0 100 100	162, 195, 247, 252	0
All	All	4111/4789 (85%)	0.07	105 (2%) 56 53	67, 119, 191, 254	0

The worst 5 of 105 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
15	S	157	TYR	7.6
1	A	69	THR	6.2
2	B	470	LYS	5.6
9	I	119	THR	5.6
2	B	250	PHE	5.0



## 6.2 Non-standard residues in protein, DNA, RNA chains [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
16	BRU	T	22	20/21	0.75	0.24	138,213,293,294	0

## 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
21	PGE	L	71	10/10	0.69	0.28	116,149,279,292	0
19	ACT	A	1734	4/4	0.81	0.42	72,95,126,151	0
19	ACT	B	1225	4/4	0.94	0.21	46,51,87,298	0
17	ZN	L	1071	1/1	0.97	0.10	158,158,158,158	0
20	EPE	A	1735	15/15	0.97	0.22	25,77,245,245	0
17	ZN	I	1122	1/1	0.97	0.04	186,186,186,186	0
17	ZN	I	1121	1/1	0.98	0.15	130,130,130,130	0
17	ZN	S	999	1/1	0.98	0.06	215,215,215,215	0
17	ZN	A	2456	1/1	0.98	0.06	145,145,145,145	0
18	MG	A	2458	1/1	0.99	0.19	83,83,83,83	0
17	ZN	B	2225	1/1	0.99	0.23	94,94,94,94	0
17	ZN	J	1066	1/1	0.99	0.24	103,103,103,103	0
17	ZN	C	1269	1/1	0.99	0.14	94,94,94,94	0
17	ZN	A	2457	1/1	0.99	0.18	87,87,87,87	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.