



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 14, 2016 – 06:10 PM EDT

PDB ID : 4BY7  
Title : elongating RNA Polymerase II-Bye1 TLD complex  
Authors : Kinkelin, K.; Wozniak, G.G.; Rothbart, S.B.; Lidschreiber, M.; Strahl, B.D.; Cramer, P.  
Deposited on : 2013-07-18  
Resolution : 3.15 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  
<http://wwpdb.org/validation/2016/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 : 1.7.1 (RC1), CSD as537be (2016)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20027790  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20027790

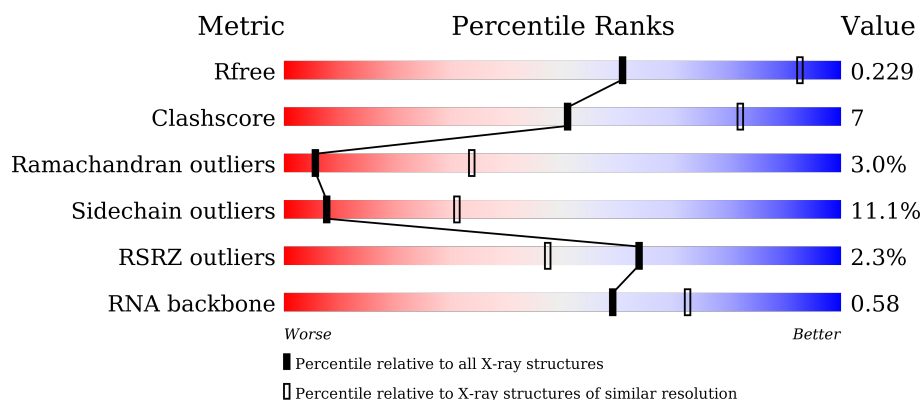
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.15 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))
$R_{free}$	91344	1112 (3.20-3.12)
Clashscore	102246	1249 (3.20-3.12)
Ramachandran outliers	100387	1222 (3.20-3.12)
Sidechain outliers	100360	1221 (3.20-3.12)
RSRZ outliers	91569	1117 (3.20-3.12)
RNA backbone	2183	1046 (3.62-2.70)

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	A	1733	<div> <div>2%</div> <div>62% 16% 18%</div> </div>
2	B	1224	<div> <div>2%</div> <div>69% 19% 9%</div> </div>
3	C	318	<div> <div>61% 20% 16%</div> </div>
4	D	221	<div> <div>61% 18% 19%</div> </div>

*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	11	
15	T	26	
16	X	146	

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
19	EPE	A	2459	-	-	-	X
19	EPE	G	1172	-	-	-	X

## 2 Entry composition

There are 19 unique types of molecules in this entry. The entry contains 33261 atoms, of which 34 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
1	A	1426	Total	C	N	O	S	0	0	0
			11223	7068	1959	2134	62			

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1115	Total	C	N	O	S	0	0	0
			8859	5609	1554	1641	55			

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3.

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

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	178	Total	C	N	O	S	0	0	0
			1387	860	246	279	2			

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	215	Total	C	N	O	S	0	0	0
			1760	1116	310	322	12			

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	87	Total	C	N	O	S	0	0	0
			705	451	119	132	3			

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7.

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

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3.

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

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9.

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

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5.

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

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	115	Total	C	N	O	S	0	0	1
			920	590	157	171	2			

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	46	Total	C	N	O	S	0	0	0
			363	224	72	63	4			

- Molecule 13 is a DNA chain called , 5'-D(\*DAP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*GP\*A

P\*GP\*CP\*DTP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	N	13	Total	C	N	O	P	0	0	0
			269	128	52	76	13			

- Molecule 14 is a RNA chain called 5'-R(\*UP\*UP\*CP\*GP\*AP\*CP\*CP\*AP\*GP\*GP\*AP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	P	10	Total	C	N	O	P	0	0	0
			215	96	41	68	10			

- Molecule 15 is a DNA chain called 5'-D(\*DAP\*GP\*CP\*TP\*CP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*DAP \*TP\*TP\*CP\*CP\*BP\*GP\*GP\*TP\*CP\*AP\*TP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	T	25	Total	Br	C	N	O	P	0	0
			504	1	240	83	155	25		

- Molecule 16 is a protein called TRANSCRIPTION FACTOR BYE1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
16	X	119	Total	C	N	O	S	0	0	0
			977	628	162	184	3			

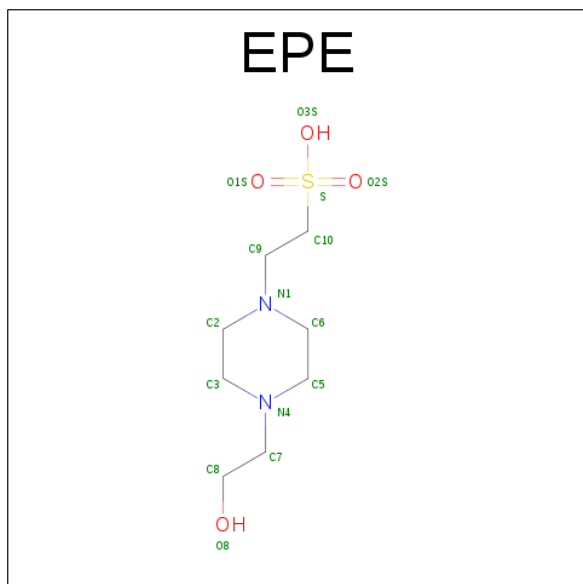
- Molecule 17 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- 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 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: C<sub>8</sub>H<sub>18</sub>N<sub>2</sub>O<sub>4</sub>S).

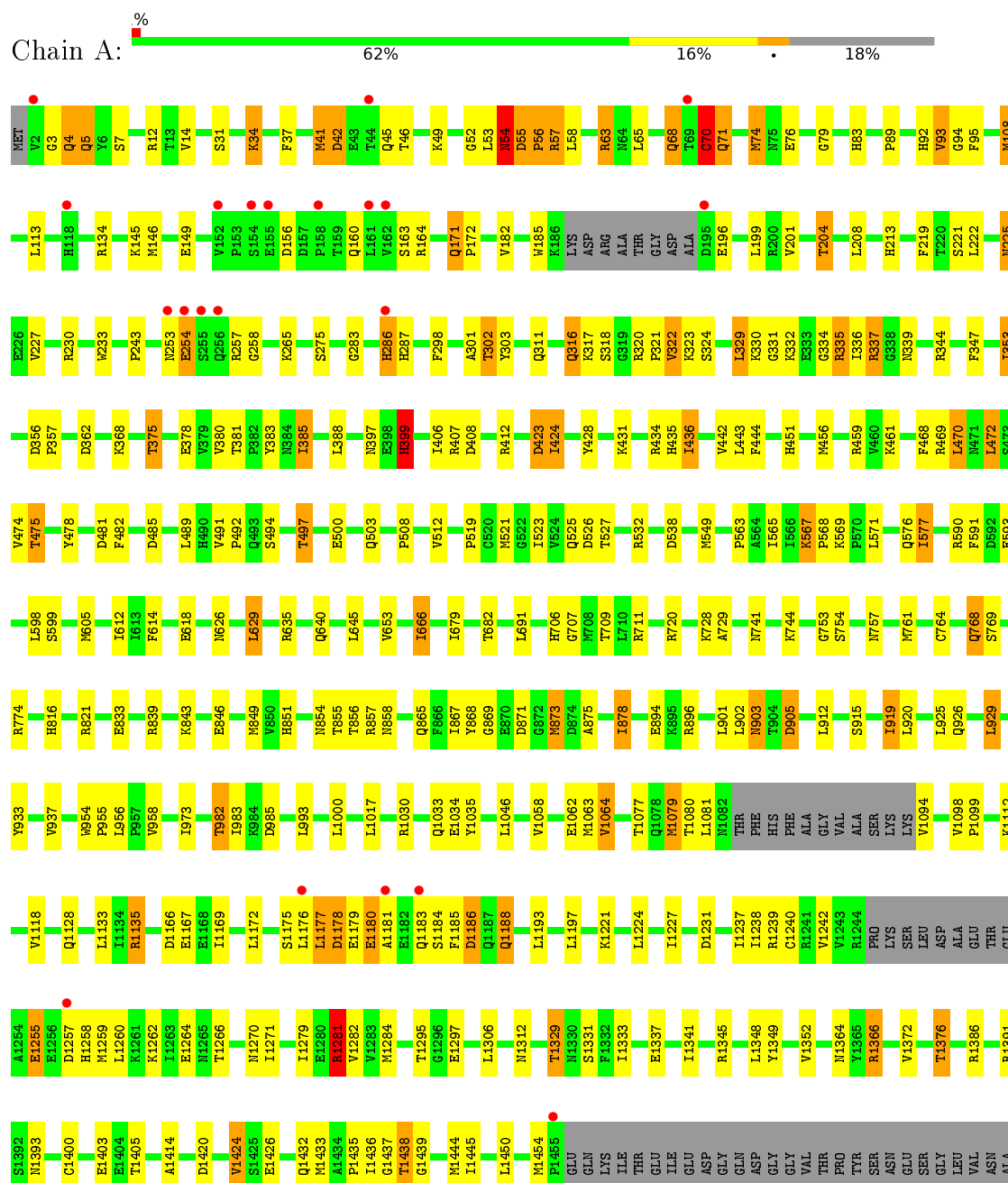


Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
19	A	1	Total	C	H	N	O	S	0	0
			32	8	17	2	4	1		
19	G	1	Total	C	H	N	O	S	0	0
			32	8	17	2	4	1		

### 3 Residue-property plots

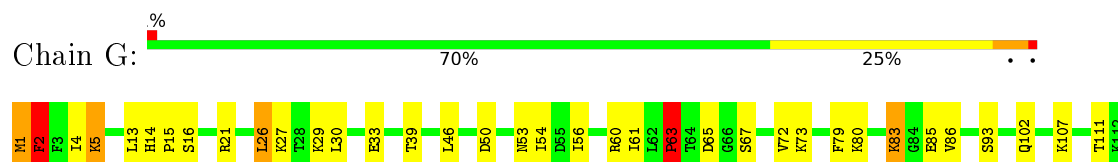
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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.

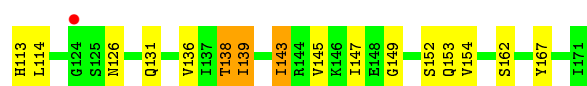
#### • Molecule 1: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB1



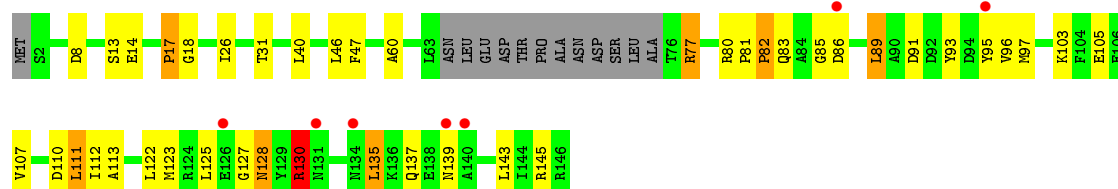




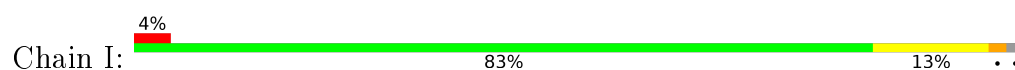




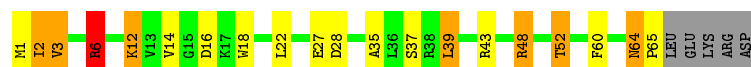
- Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3



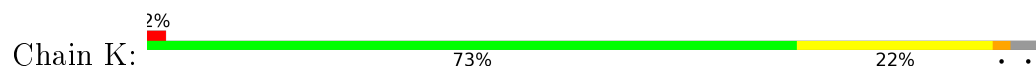
- Molecule 9: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9



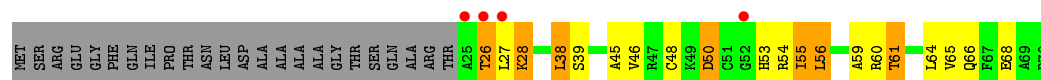
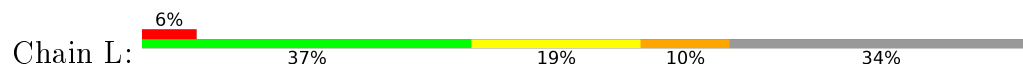
- Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5



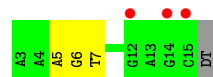
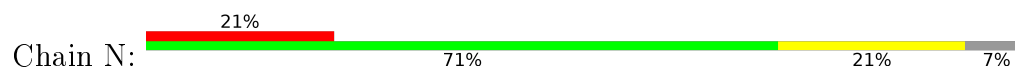
- Molecule 11: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11



- Molecule 12: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4



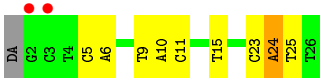
- Molecule 13: , 5'-D(\*DAP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*GP\*AP\*GP\*CP\*DTP)-3'



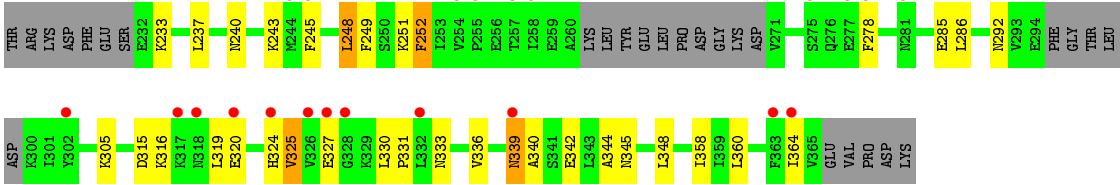
- Molecule 14: 5'-R(\*UP\*UP\*CP\*GP\*AP\*CP\*CP\*AP\*GP\*GP\*AP)-3'



• Molecule 15: 5'-D(\*DAP\*GP\*CP\*TP\*CP\*AP\*AP\*GP\*TP\*AP\*CP\*TP\*TP\*DAP \*TP\*TP\*CP\*CP\*BP\*GP\*GP\*TP\*CP\*AP\*TP\*T)-3'



• Molecule 16: TRANSCRIPTION FACTOR BYE1



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	222.50Å 390.68Å 281.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.84 – 3.15 48.84 – 3.15	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.84-3.15) 100.0 (48.84-3.15)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.60 (at 3.12Å)	Xtriage
Refinement program	BUSTER 2.11.5	Depositor
R, $R_{free}$	0.189 , 0.212 0.206 , 0.229	Depositor DCC
$R_{free}$ test set	4003 reflections (1.94%)	DCC
Wilson B-factor (Å <sup>2</sup> )	99.4	Xtriage
Anisotropy	0.406	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 83.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.003 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.000 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	33261	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	115.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.83% 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: ZN, EPE, MG, BRU

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	0.51	1/11424 (0.0%)	0.78	5/15452 (0.0%)
2	B	0.48	0/9030	0.74	4/12174 (0.0%)
3	C	0.46	0/2133	0.74	0/2891
4	D	0.48	0/1397	0.75	0/1878
5	E	0.43	0/1796	0.63	0/2416
6	F	0.55	0/717	0.78	1/967 (0.1%)
7	G	0.47	0/1368	0.78	2/1844 (0.1%)
8	H	0.46	0/1085	0.72	0/1467
9	I	0.45	0/989	0.70	0/1331
10	J	0.49	0/541	0.82	0/727
11	K	0.42	0/938	0.64	0/1267
12	L	0.53	0/364	0.85	0/482
13	N	1.05	0/302	1.03	0/464
14	P	1.10	1/240 (0.4%)	0.94	0/372
15	T	1.22	1/539 (0.2%)	1.11	2/825 (0.2%)
16	X	0.47	0/990	0.70	0/1325
All	All	0.52	3/33853 (0.0%)	0.76	14/45882 (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
15	T	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	T	24	DA	C3'-O3'	6.30	1.52	1.44
1	A	56	PRO	C-N	5.51	1.46	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	P	3	C	C1'-N1	5.49	1.56	1.48

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	1177	LEU	C-N-CA	6.51	137.99	121.70
15	T	24	DA	O4'-C1'-N9	6.48	112.53	108.00
7	G	1	MET	C-N-CA	6.24	137.31	121.70
1	A	399	HIS	N-CA-CB	6.17	121.72	110.60
1	A	56	PRO	C-N-CA	6.05	136.83	121.70

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
15	T	15	DT	Sidechain
15	T	25	DT	Sidechain

## 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	A	11223	0	11273	187	0
2	B	8859	0	8902	118	0
3	C	2095	0	2051	36	0
4	D	1387	0	1366	20	0
5	E	1760	0	1788	22	0
6	F	705	0	731	14	0
7	G	1340	0	1357	33	0
8	H	1068	0	1039	30	0
9	I	971	0	927	6	0
10	J	532	0	542	11	0
11	K	920	0	929	15	0
12	L	363	0	385	7	0
13	N	269	0	147	3	0
14	P	215	0	109	1	0
15	T	504	0	274	4	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
16	X	977	0	996	13	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
17	J	1	0	0	0	0
17	L	1	0	0	0	0
18	A	1	0	0	0	0
19	A	15	17	17	1	0
19	G	15	17	17	0	0
All	All	33227	34	32850	458	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.

The worst 5 of 458 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:53:LEU:HD23	1:A:54:ASN:H	1.23	1.02
1:A:567:LYS:HB2	1:A:568:PRO:HD3	1.45	0.94
1:A:567:LYS:HD2	1:A:568:PRO:HD3	1.49	0.94
1:A:1364:ASN:ND2	1:A:1366:ARG:HD2	1.83	0.93
4:D:40:HIS:HB3	7:G:73:LYS:HE3	1.52	0.92

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%)	1274 (90%)	96 (7%)	48 (3%)	5	29
2	B	1097/1224 (90%)	981 (89%)	84 (8%)	32 (3%)	6	35

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	264/318 (83%)	246 (93%)	14 (5%)	4 (2%)	13	53
4	D	174/221 (79%)	156 (90%)	11 (6%)	7 (4%)	4	25
5	E	213/215 (99%)	209 (98%)	4 (2%)	0	100	100
6	F	85/155 (55%)	79 (93%)	4 (5%)	2 (2%)	7	41
7	G	169/171 (99%)	153 (90%)	12 (7%)	4 (2%)	7	41
8	H	128/146 (88%)	103 (80%)	15 (12%)	10 (8%)	1	7
9	I	117/122 (96%)	102 (87%)	13 (11%)	2 (2%)	11	50
10	J	63/70 (90%)	55 (87%)	5 (8%)	3 (5%)	3	20
11	K	113/120 (94%)	112 (99%)	1 (1%)	0	100	100
12	L	42/70 (60%)	32 (76%)	5 (12%)	5 (12%)	0	2
16	X	113/146 (77%)	101 (89%)	8 (7%)	4 (4%)	4	29
All	All	3996/4711 (85%)	3603 (90%)	272 (7%)	121 (3%)	5	34

5 of 121 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	54	ASN
1	A	57	ARG
1	A	76	GLU
1	A	253	ASN
1	A	257	ARG

### 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	
1	A	1248/1520 (82%)	1115 (89%)	133 (11%)	8	32
2	B	966/1061 (91%)	860 (89%)	106 (11%)	8	31
3	C	234/274 (85%)	212 (91%)	22 (9%)	11	39
4	D	147/200 (74%)	132 (90%)	15 (10%)	9	35
5	E	197/197 (100%)	186 (94%)	11 (6%)	26	66

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
6	F	77/137 (56%)	66 (86%)	11 (14%)	4	19
7	G	152/152 (100%)	133 (88%)	19 (12%)	6	25
8	H	117/128 (91%)	100 (86%)	17 (14%)	4	18
9	I	113/116 (97%)	103 (91%)	10 (9%)	12	43
10	J	60/65 (92%)	48 (80%)	12 (20%)	1	8
11	K	99/102 (97%)	89 (90%)	10 (10%)	9	35
12	L	40/57 (70%)	28 (70%)	12 (30%)	0	1
16	X	109/136 (80%)	93 (85%)	16 (15%)	4	18
All	All	3559/4145 (86%)	3165 (89%)	394 (11%)	8	31

5 of 394 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	563	MET
2	B	1045	SER
11	K	74	ARG
2	B	617	ARG
2	B	815	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 80 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	366	GLN
2	B	862	GLN
9	I	12	ASN
2	B	449	ASN
2	B	590	HIS

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
14	P	9/11 (81%)	3 (33%)	0

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
14	P	3	C

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
14	P	5	A
14	P	9	G

There are no RNA pucker outliers to report.

## 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	$\# Z  > 2$
15	BRU	T	19	15,14	12,21,22	2.36	1 (8%)	16,30,33	2.73	3 (18%)

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
15	BRU	T	19	15,14	-	0/3/21/22	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
15	T	19	BRU	C4-C5	7.89	1.48	1.38

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
15	T	19	BRU	C5-C4-N3	-5.26	118.38	124.00
15	T	19	BRU	O4'-C1'-N1	2.77	112.55	107.71
15	T	19	BRU	C4-N3-C2	9.06	122.72	115.16

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 11 ligands modelled in this entry, 9 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	$\# Z  > 2$
19	EPE	A	2459	-	15,15,15	0.94	1 (6%)	19,20,20	0.65	0
19	EPE	G	1172	-	15,15,15	1.14	1 (6%)	19,20,20	0.52	0

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
19	EPE	A	2459	-	-	0/9/19/19	0/1/1/1
19	EPE	G	1172	-	-	0/9/19/19	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
19	G	1172	EPE	C10-S	-4.35	1.71	1.77
19	A	2459	EPE	C10-S	-3.54	1.72	1.77

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
19	A	2459	EPE	1	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

### 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	A	1426/1733 (82%)	-0.07	21 (1%) 76 62	60, 98, 157, 213	0
2	B	1115/1224 (91%)	0.04	20 (1%) 71 56	61, 107, 165, 198	0
3	C	266/318 (83%)	-0.16	0 100 100	74, 96, 135, 164	0
4	D	178/221 (80%)	-0.15	1 (0%) 90 84	79, 111, 157, 183	0
5	E	215/215 (100%)	0.06	7 (3%) 50 33	80, 134, 176, 192	0
6	F	87/155 (56%)	-0.24	0 100 100	67, 81, 116, 141	0
7	G	171/171 (100%)	-0.07	1 (0%) 90 84	77, 99, 135, 155	0
8	H	133/146 (91%)	0.31	7 (5%) 30 16	103, 140, 180, 198	0
9	I	119/122 (97%)	0.12	5 (4%) 40 24	108, 139, 175, 194	0
10	J	65/70 (92%)	-0.19	0 100 100	82, 96, 138, 151	0
11	K	115/120 (95%)	-0.04	2 (1%) 73 59	68, 98, 135, 156	0
12	L	46/70 (65%)	1.06	4 (8%) 13 6	77, 172, 184, 191	0
13	N	13/14 (92%)	0.84	3 (23%) 1 1	150, 182, 258, 261	0
14	P	10/11 (90%)	-0.34	0 100 100	85, 108, 157, 162	0
15	T	24/26 (92%)	0.18	2 (8%) 14 6	98, 147, 265, 267	0
16	X	119/146 (81%)	0.82	22 (18%) 2 1	160, 190, 211, 222	0
All	All	4102/4762 (86%)	0.01	95 (2%) 64 47	60, 106, 177, 267	0

The worst 5 of 95 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
12	L	25	ALA	8.7
11	K	115	ALA	8.4
12	L	26	THR	7.9
2	B	250	PHE	6.1
1	A	69	THR	5.7

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
15	BRU	T	19	20/21	0.96	0.13	-	91,98,105,114	0

## 6.3 Carbohydrates [i](#)

There are no carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
19	EPE	A	2459	15/15	0.24	1.15	21.37	249,261,272,273	0
19	EPE	G	1172	15/15	0.37	0.85	12.47	217,230,230,231	0
17	ZN	B	2225	1/1	1.00	0.22	0.54	87,87,87,87	0
17	ZN	I	1121	1/1	0.99	0.15	0.33	117,117,117,117	0
17	ZN	A	2457	1/1	0.99	0.19	-0.45	80,80,80,80	0
17	ZN	J	1066	1/1	1.00	0.22	-0.97	84,84,84,84	0
17	ZN	C	1269	1/1	1.00	0.11	-1.14	88,88,88,88	0
17	ZN	A	2456	1/1	0.99	0.08	-1.61	124,124,124,124	0
17	ZN	I	1122	1/1	0.95	0.04	-2.07	191,191,191,191	0
17	ZN	L	1071	1/1	0.97	0.04	-5.65	193,193,193,193	0
18	MG	A	2458	1/1	0.92	0.12	-	275,275,275,275	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.