wwPDB X-ray Structure Validation Summary Report (i)

Sep 23, 2024 – 02:17 PM EDT

PDB ID	:	9BVT
Title	:	RNA Pol II - High Mn(+2) concentration
Authors	:	Calero, G.
Deposited on	:	2024-05-20
Resolution	:	3.40 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

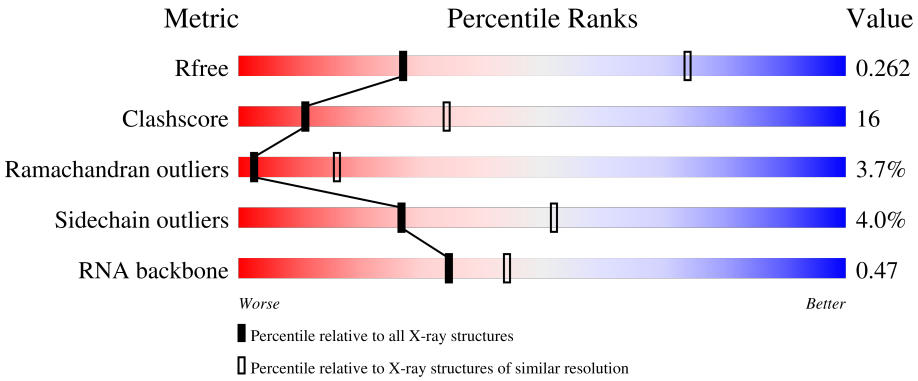
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.002 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.3

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.40 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.



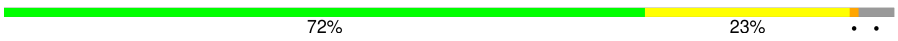
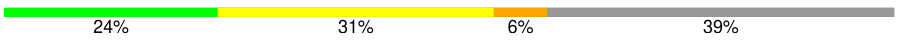
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	1140 (3.46-3.34)
Clashscore	180529	1172 (3.46-3.34)
Ramachandran outliers	177936	1172 (3.46-3.34)
Sidechain outliers	177891	1172 (3.46-3.34)
RNA backbone	3690	1033 (3.80-3.00)

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 >=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 <=5%

Mol	Chain	Length	Quality of chain
1	A	1733	48% 29% • 20%
2	B	1224	52% 32% • 13%
3	C	318	55% 26% • 16%
4	D	221	47% 25% • 27%
5	E	215	63% 31% • •
6	F	155	32% 21% • 46%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
7	G	171	
8	H	146	
9	I	122	
10	J	70	
11	K	120	
12	L	70	
13	X	10	
14	W	13	

2 Entry composition (i)

There are 18 unique types of molecules in this entry. The entry contains 30742 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	1378	10848	6846	1894	2047	61	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	1060	8428	5349	1476	1549	54	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	162	1287	799	224	262	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	208	1713	1089	303	312	9	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
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 RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	117	Total	C	N	O	S	0	0	0
			951	605	158	184	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 II subunit RPABC5.

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 II subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	43	Total	C	N	O	S	0	0	0
			343	211	69	59	4			

- Molecule 13 is a RNA chain called RNA (5'-R(*AP*UP*CP*GP*AP*GP*AP*GP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	X	10	Total	C	N	O	P	0	0	0
			217	98	45	65	9			

- Molecule 14 is a DNA chain called DNA (5'-D(P*AP*CP*GP*TP*CP*CP*CP*TP*CP*TP*CP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	W	13	Total	C	N	O	P	0	0	0
			260	124	44	79	13			

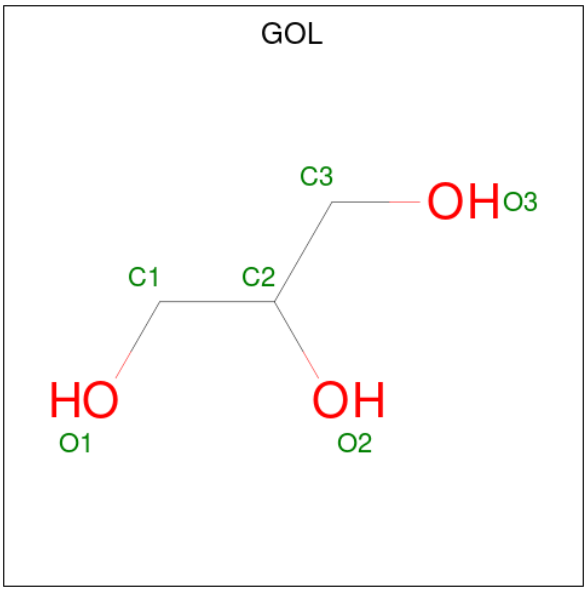
- Molecule 15 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 16 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	A	2	Total	Mn	0	0
			2	2		

- Molecule 17 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
17	E	1	Total C O 6 3 3	0	0

- Molecule 18 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
18	A	55	Total O 55 55	0	0
18	B	45	Total O 45 45	0	0
18	C	7	Total O 7 7	0	0
18	D	5	Total O 5 5	0	0
18	E	6	Total O 6 6	0	0
18	F	4	Total O 4 4	0	0
18	G	4	Total O 4 4	0	0
18	H	2	Total O 2 2	0	0
18	J	1	Total O 1 1	0	0
18	K	6	Total O 6 6	0	0
18	L	3	Total O 3 3	0	0

Continued on next page...

Continued from previous page...

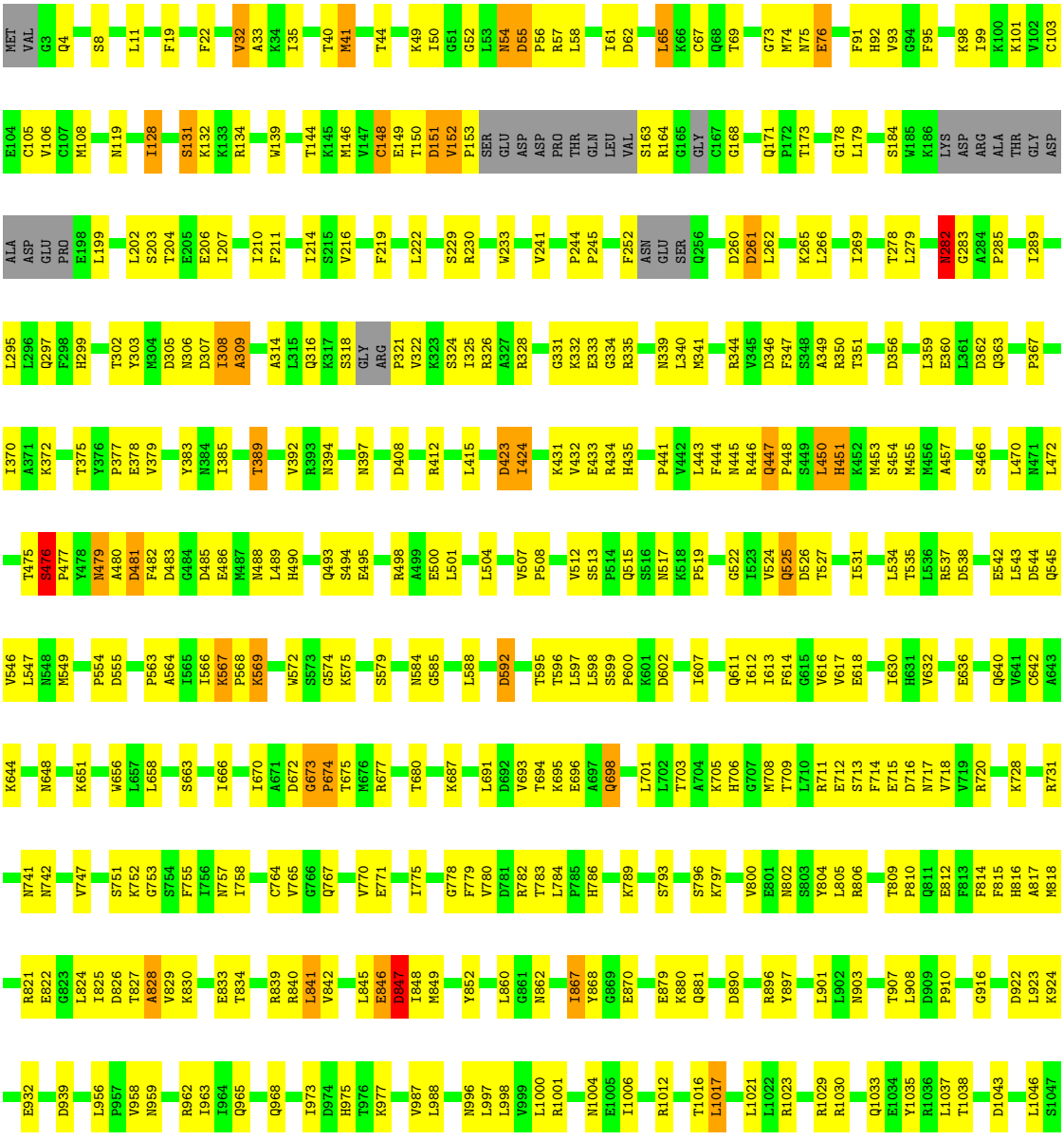
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
18	X	2	Total O 2 2	0	0
18	W	2	Total O 2 2	0	0

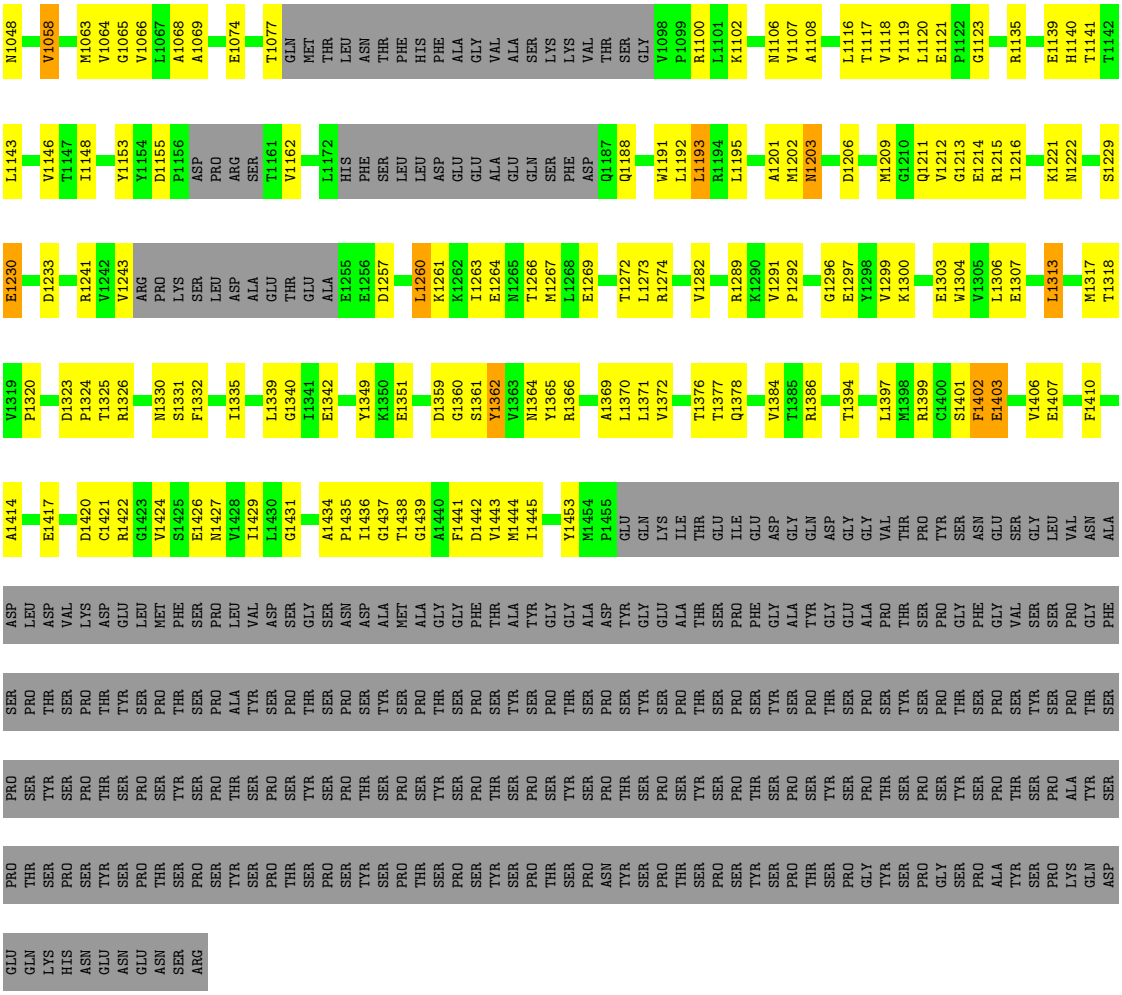
3 Residue-property plots (i)

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

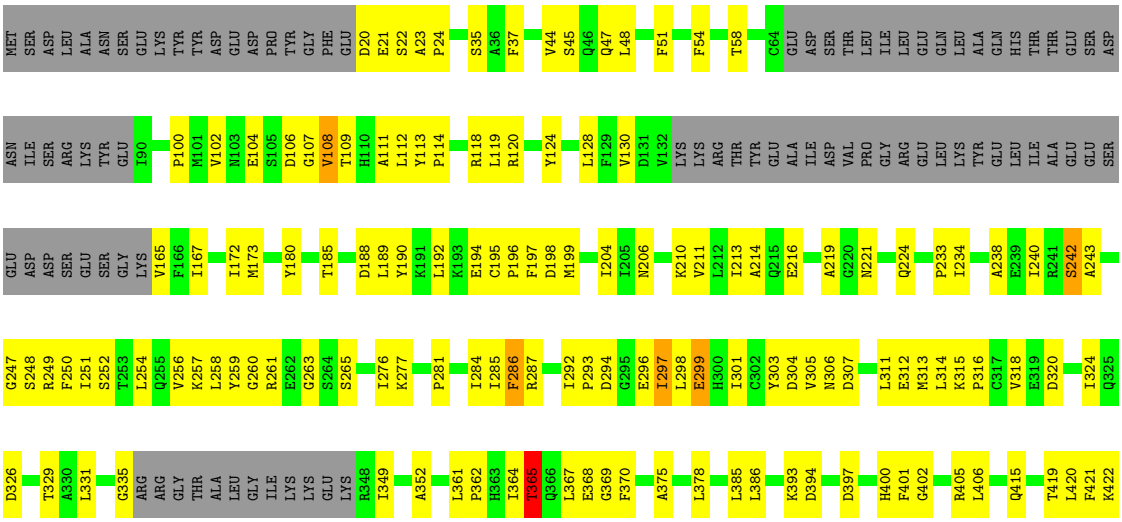
Chain A:

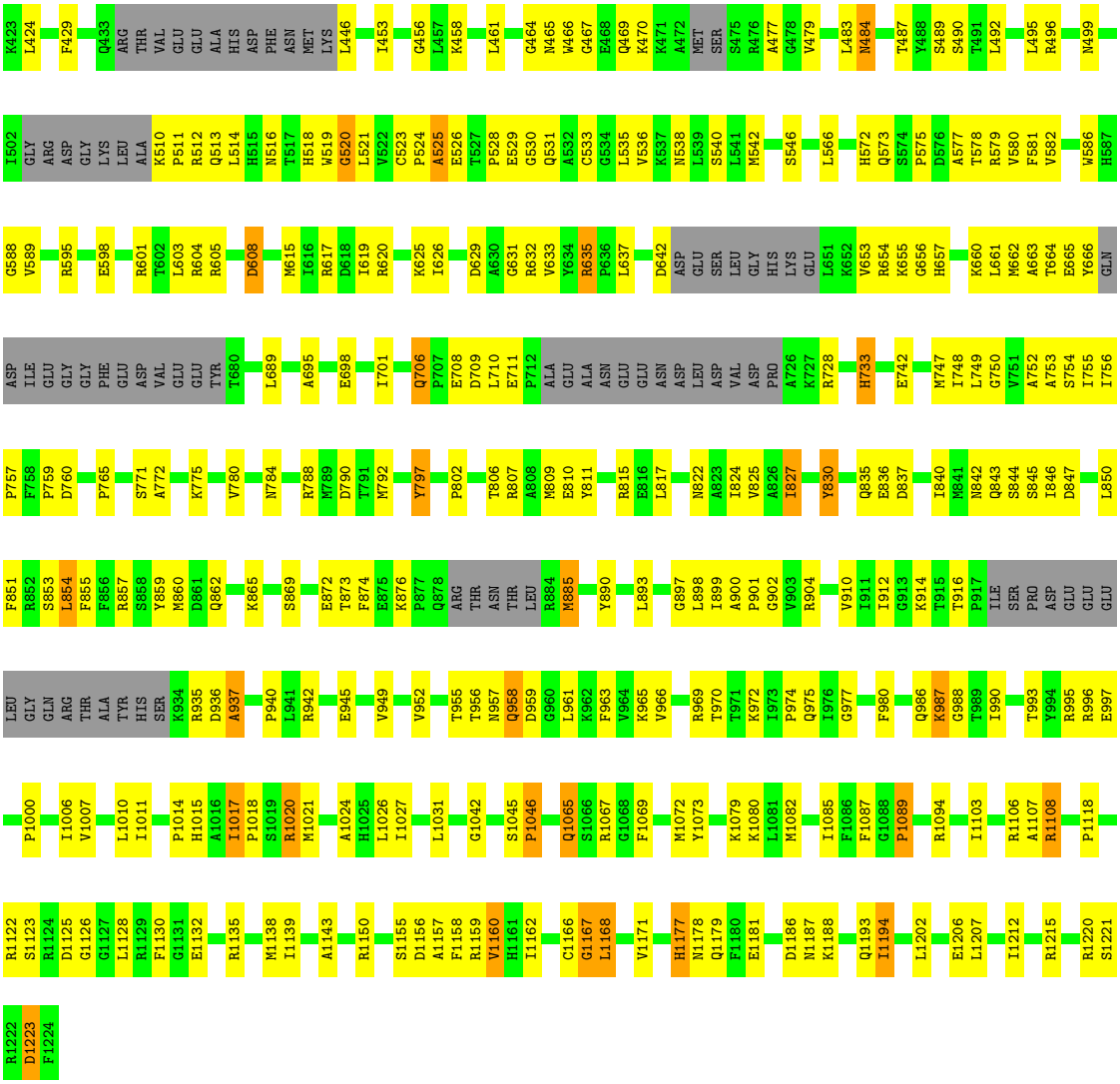




• Molecule 2: DNA-directed RNA polymerase subunit beta

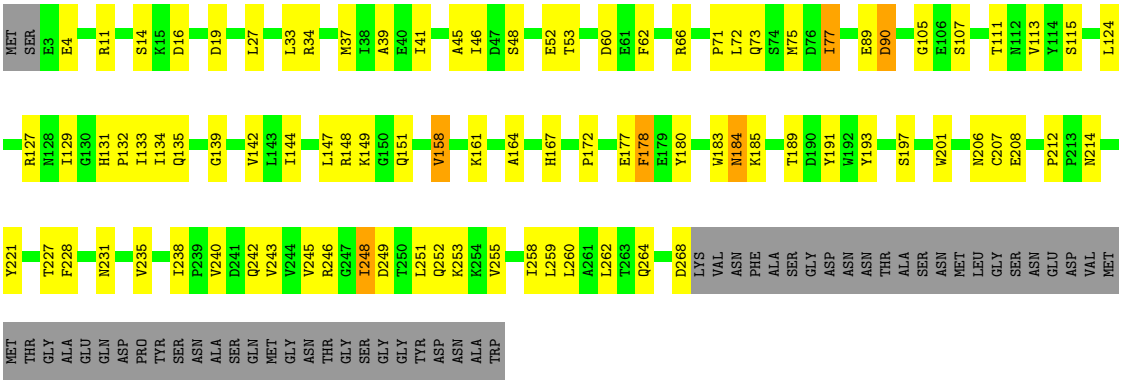
Chain B:





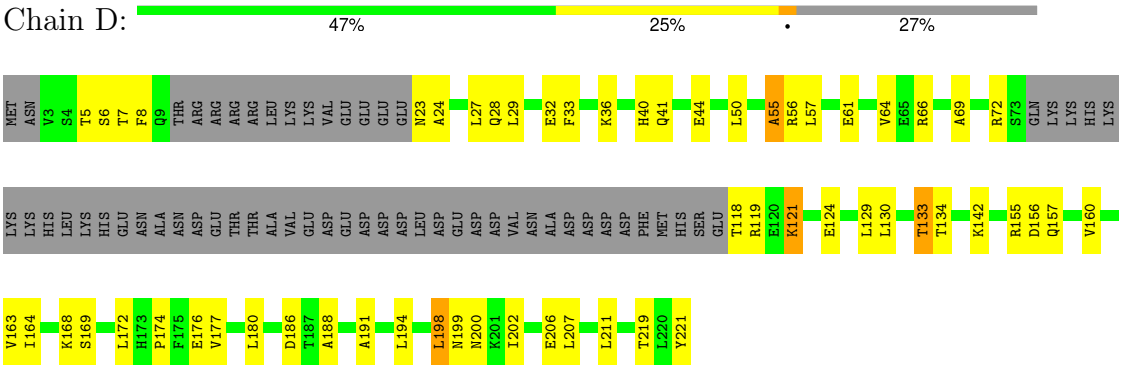
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3

Chain C:



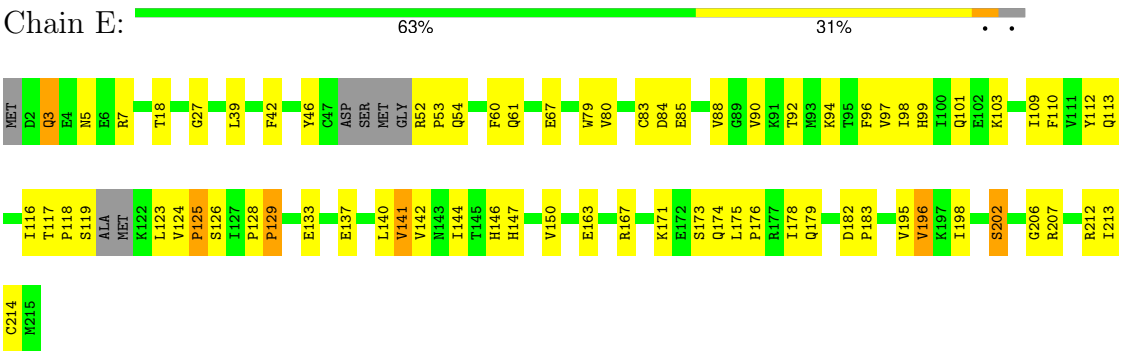
• Molecule 4: DNA-directed RNA polymerase II subunit RPB4

Chain D:



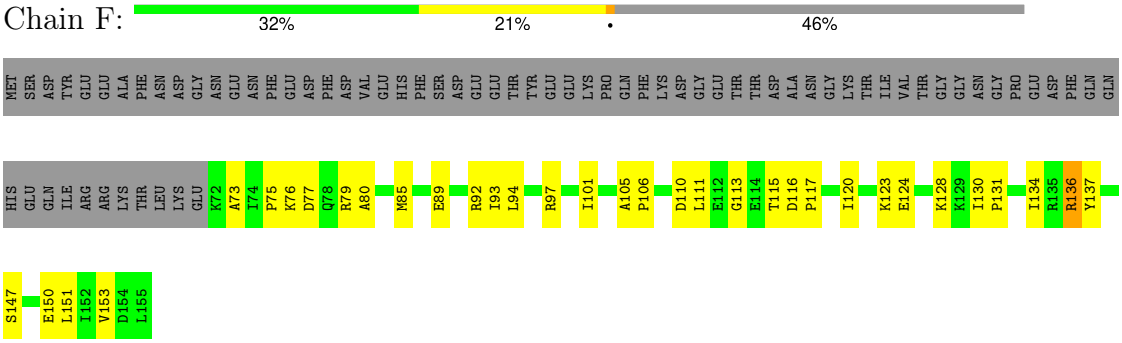
• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E:



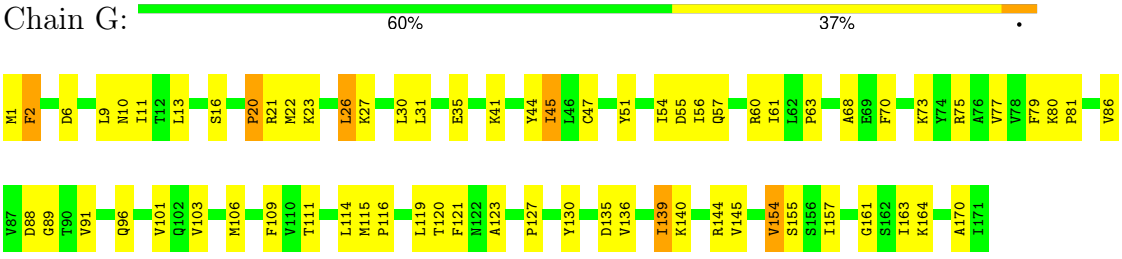
• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F:



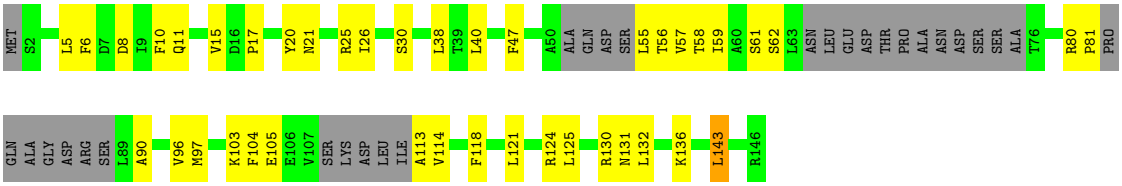
• Molecule 7: DNA-directed RNA polymerase II subunit RPB7

Chain G:



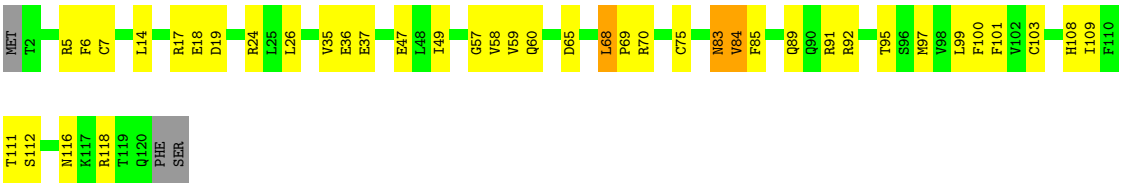
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:



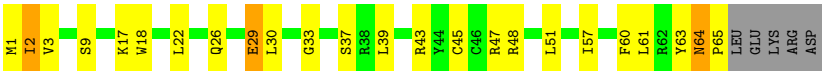
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9

Chain I:



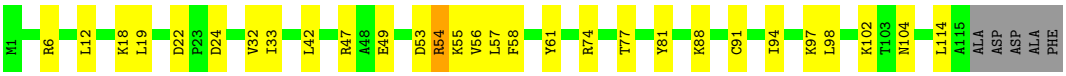
• Molecule 10: DNA-directed RNA polymerases II subunit RPABC5

Chain J:



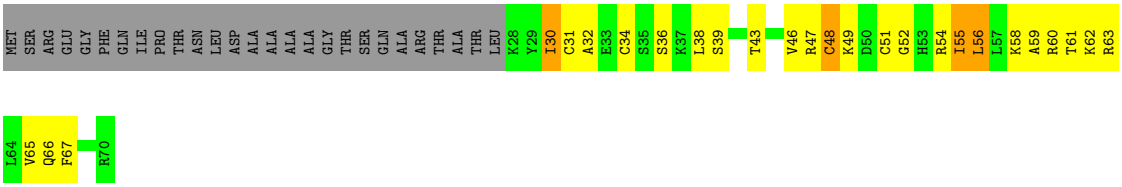
• Molecule 11: DNA-directed RNA polymerase II subunit RPB11

Chain K:



• Molecule 12: DNA-directed RNA polymerases II subunit RPABC4

Chain L:



• Molecule 13: RNA (5'-R(*AP*UP*CP*GP*AP*GP*AP*GP*GP*A)-3')

Chain X:

• Molecule 14: DNA (5'-D(P*AP*CP*GP*TP*CP*CP*CP*TP*CP*TP*CP*GP*A)-3')

Chain W:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	221.19Å 393.97Å 283.28Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.25 – 3.40 49.25 – 3.40	Depositor EDS
% Data completeness (in resolution range)	89.0 (49.25-3.40) 98.6 (49.25-3.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.02 (at 3.33Å)	Xtriage
Refinement program	PHENIX (1.19.2_4158: ???)	Depositor
R, R_{free}	0.220 , 0.249 0.242 , 0.262	Depositor DCC
R_{free} test set	8254 reflections (4.89%)	wwPDB-VP
Wilson B-factor (Å^2)	92.3	Xtriage
Anisotropy	1.597	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 145.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.40$, $< L^2 > = 0.23$	Xtriage
Estimated twinning fraction	0.066 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.076 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	30742	wwPDB-VP
Average B, all atoms (Å^2)	157.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.10% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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, MN, GOL

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	0.29	0/11037	0.54	0/14916
2	B	0.28	0/8593	0.55	0/11585
3	C	0.25	0/2133	0.49	0/2891
4	D	0.25	0/1296	0.49	0/1741
5	E	0.26	0/1747	0.53	0/2349
6	F	0.24	0/691	0.51	0/933
7	G	0.26	0/1368	0.49	0/1844
8	H	0.25	0/965	0.55	0/1302
9	I	0.25	0/989	0.57	0/1331
10	J	0.27	0/541	0.55	0/727
11	K	0.25	0/938	0.48	0/1267
12	L	0.31	0/345	0.59	0/457
13	X	0.25	0/244	0.72	0/380
14	W	0.62	0/289	1.12	0/442
All	All	0.28	0/31176	0.54	0/42165

There are no bond length outliers.

There are no bond angle outliers.

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	A	10848	0	10932	439	0
2	B	8428	0	8465	316	0
3	C	2095	0	2051	76	0
4	D	1287	0	1296	28	0
5	E	1713	0	1739	56	0
6	F	679	0	701	26	0
7	G	1340	0	1357	42	0
8	H	951	0	926	32	0
9	I	971	0	929	24	0
10	J	532	0	542	21	0
11	K	920	0	929	22	0
12	L	343	0	363	23	0
13	X	217	0	110	10	0
14	W	260	0	147	12	0
15	A	2	0	0	0	0
15	B	1	0	0	0	0
15	C	1	0	0	0	0
15	I	2	0	0	0	0
15	J	1	0	0	0	0
15	L	1	0	0	0	0
16	A	2	0	0	0	0
17	E	6	0	8	1	0
18	A	55	0	0	2	0
18	B	45	0	0	1	0
18	C	7	0	0	0	0
18	D	5	0	0	0	0
18	E	6	0	0	0	0
18	F	4	0	0	0	0
18	G	4	0	0	0	0
18	H	2	0	0	0	0
18	J	1	0	0	0	0
18	K	6	0	0	0	0
18	L	3	0	0	0	0
18	W	2	0	0	0	0
18	X	2	0	0	0	0
All	All	30742	0	30495	989	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 16.

The worst 5 of 989 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)
13:X:10:A:N6	14:W:20:DC:H42	1.49	1.09
1:A:33:ALA:HB3	1:A:56:PRO:O	1.64	0.97
2:B:952:VAL:HG23	2:B:966:VAL:HG22	1.55	0.89
13:X:10:A:N6	14:W:20:DC:N4	2.21	0.87
1:A:56:PRO:HD2	1:A:58:LEU:HB2	1.56	0.85

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	1358/1733 (78%)	1135 (84%)	169 (12%)	54 (4%)	2 15
2	B	1036/1224 (85%)	869 (84%)	129 (12%)	38 (4%)	2 16
3	C	264/318 (83%)	239 (90%)	19 (7%)	6 (2%)	5 23
4	D	156/221 (71%)	133 (85%)	15 (10%)	8 (5%)	1 11
5	E	202/215 (94%)	181 (90%)	17 (8%)	4 (2%)	6 25
6	F	82/155 (53%)	74 (90%)	7 (8%)	1 (1%)	11 35
7	G	169/171 (99%)	141 (83%)	19 (11%)	9 (5%)	1 10
8	H	107/146 (73%)	84 (78%)	19 (18%)	4 (4%)	2 16
9	I	117/122 (96%)	99 (85%)	12 (10%)	6 (5%)	1 11
10	J	63/70 (90%)	53 (84%)	8 (13%)	2 (3%)	3 18
11	K	113/120 (94%)	102 (90%)	10 (9%)	1 (1%)	14 41
12	L	41/70 (59%)	21 (51%)	15 (37%)	5 (12%)	0 2
All	All	3708/4565 (81%)	3131 (84%)	439 (12%)	138 (4%)	2 16

5 of 138 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	32	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	76	GLU
1	A	119	ASN
1	A	131	SER
1	A	424	ILE

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	1204/1520 (79%)	1156 (96%)	48 (4%)	27 52
2	B	918/1061 (86%)	879 (96%)	39 (4%)	25 51
3	C	234/274 (85%)	229 (98%)	5 (2%)	48 69
4	D	144/200 (72%)	134 (93%)	10 (7%)	13 38
5	E	192/197 (98%)	187 (97%)	5 (3%)	41 64
6	F	74/137 (54%)	73 (99%)	1 (1%)	62 77
7	G	152/152 (100%)	146 (96%)	6 (4%)	27 53
8	H	104/128 (81%)	100 (96%)	4 (4%)	28 54
9	I	113/116 (97%)	108 (96%)	5 (4%)	24 50
10	J	60/65 (92%)	57 (95%)	3 (5%)	20 46
11	K	99/102 (97%)	97 (98%)	2 (2%)	50 70
12	L	38/57 (67%)	34 (90%)	4 (10%)	5 21
All	All	3332/4009 (83%)	3200 (96%)	132 (4%)	27 52

5 of 132 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
8	H	56	THR
9	I	68	LEU
12	L	55	ILE
2	B	45	SER
1	A	1453	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	538	ASN
2	B	986	GLN
11	K	40	HIS
3	C	135	GLN
7	G	57	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
13	X	9/10 (90%)	2 (22%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
13	X	8	G
13	X	10	A

There are no RNA pucker outliers to report.

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 11 ligands modelled in this entry, 10 are monoatomic - leaving 1 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	GOL	E	301	-	5,5,5	0.96	0	5,5,5	1.06	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
17	GOL	E	301	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
17	E	301	GOL	O1-C1-C2-O2
17	E	301	GOL	O1-C1-C2-C3
17	E	301	GOL	C1-C2-C3-O3
17	E	301	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
17	E	301	GOL	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

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.