



wwPDB EM Model Validation Summary Report ⓘ

Mar 15, 2020 – 03:42 PM EDT

PDB ID : 6PSR
EMDB ID : EMD-20461
Title : Escherichia coli RNA polymerase promoter unwinding intermediate (TRPi1)
with TraR and rpsT P2 promoter
Authors : Chen, J.; Chiu, C.E.; Campbell, E.A.; Darst, S.A.
Deposited on : 2019-07-13
Resolution : 3.40 Å(reported)

This is a wwPDB EM Model Validation Summary Report for a publicly released PDB/EMDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/EMValidationReportHelp>

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:

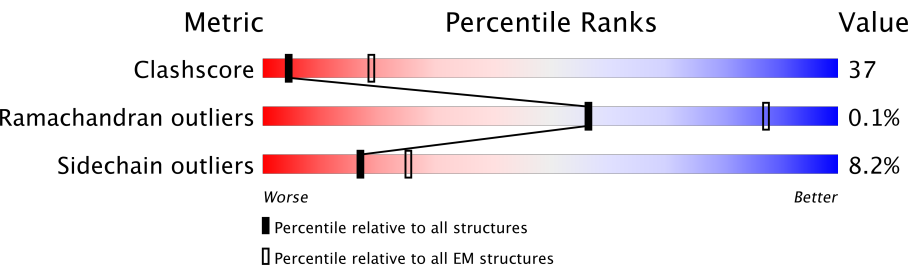
MolProbity : 4.02b-467
Mogul : 1.8.0 (224370), CSD as540be (2019)
buster-report : 1.1.7 (2018)
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.8

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

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)	EM structures (#Entries)
Clashscore	136327	1886
Ramachandran outliers	132723	1663
Sidechain outliers	132532	1531

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments on the 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	G	329	35% 32% . 30%
1	H	329	27% 37% . 33%
1	M	329	5% 16% . 78%
2	I	1342	42% 52% 5%
3	J	1430	39% 51% . 6%
4	K	91	33% 44% 5% 18%
5	L	616	34% 53% . 11%
6	N	72	54% 40% 6%
7	O	85	9% 33% 58%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	P	85	<div><div></div><div>39%</div><div>59%</div></div>

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 32342 atoms, of which 156 are hydrogens and 0 are deuteriums.

In the tables below, 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 subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
1	G	230	Total	C	N	O	S	0	0
			1771	1106	314	345	6		
1	H	219	Total	C	N	O	S	0	0
			1678	1048	295	329	6		
1	M	73	Total	C	N	O	S	0	0
			572	362	100	108	2		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	I	1340	Total	C	N	O	S	0	0
			10559	6627	1841	2048	43		

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					AltConf	Trace
3	J	1345	Total	C	N	O	S	0	0
			10466	6577	1867	1972	50		

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	1	VAL	-	expression tag	UNP P0A8T7
J	1408	LEU	-	expression tag	UNP P0A8T7
J	1409	GLU	-	expression tag	UNP P0A8T7
J	1410	LEU	-	expression tag	UNP P0A8T7
J	1411	GLU	-	expression tag	UNP P0A8T7
J	1412	VAL	-	expression tag	UNP P0A8T7
J	1413	LEU	-	expression tag	UNP P0A8T7
J	1414	PHE	-	expression tag	UNP P0A8T7
J	1415	GLN	-	expression tag	UNP P0A8T7
J	1416	GLY	-	expression tag	UNP P0A8T7
J	1417	PRO	-	expression tag	UNP P0A8T7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
J	1418	SER	-	expression tag	UNP P0A8T7
J	1419	SER	-	expression tag	UNP P0A8T7
J	1420	GLY	-	expression tag	UNP P0A8T7
J	1421	HIS	-	expression tag	UNP P0A8T7
J	1422	HIS	-	expression tag	UNP P0A8T7
J	1423	HIS	-	expression tag	UNP P0A8T7
J	1424	HIS	-	expression tag	UNP P0A8T7
J	1425	HIS	-	expression tag	UNP P0A8T7
J	1426	HIS	-	expression tag	UNP P0A8T7
J	1427	HIS	-	expression tag	UNP P0A8T7
J	1428	HIS	-	expression tag	UNP P0A8T7
J	1429	HIS	-	expression tag	UNP P0A8T7
J	1430	HIS	-	expression tag	UNP P0A8T7

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

Mol	Chain	Residues	Atoms					AltConf	Trace
4	K	75	Total	C	N	O	S	0	0
			600	365	114	120	1		

- Molecule 5 is a protein called RNA polymerase sigma factor RpoD.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	L	548	Total	C	N	O	S	0	0
			4407	2754	771	855	27		

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
L	-2	SER	-	expression tag	UNP Q0P6L9
L	-1	GLU	-	expression tag	UNP Q0P6L9
L	0	PHE	-	expression tag	UNP Q0P6L9

- Molecule 6 is a protein called Protein TraR.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	N	72	Total	C	N	O	S	0	0
			565	350	102	108	5		

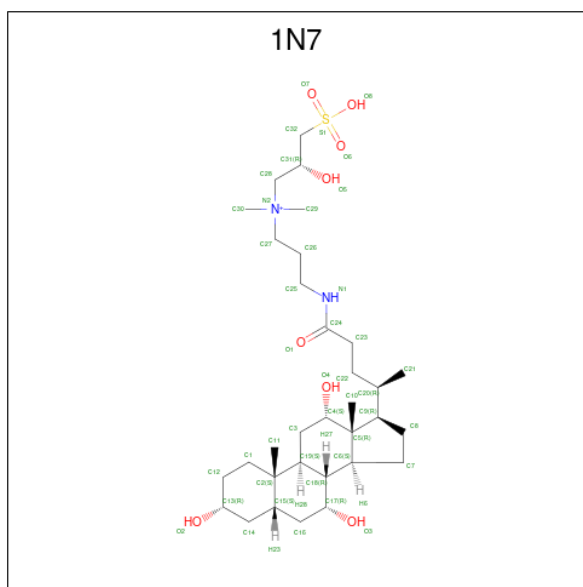
- Molecule 7 is a DNA chain called DNA (85-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
7	O	36	Total	C	N	O	P	0	0
			746	353	154	203	36		

- Molecule 8 is a DNA chain called DNA (85-MER).

Mol	Chain	Residues	Atoms					AltConf	Trace
8	P	35	Total	C	N	O	P	0	0
			710	342	111	222	35		

- Molecule 9 is CHAPSO (three-letter code: 1N7) (formula: C₃₂H₅₉N₂O₈S).



Mol	Chain	Residues	Atoms				AltConf
9	I	1	Total	C	H	O	0
			66	24	39	3	
9	J	1	Total	C	H	O	0
			132	48	78	6	
9	J	1	Total	C	H	O	0
			132	48	78	6	
9	L	1	Total	C	H	O	0
			66	24	39	3	

- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		AltConf
10	J	1	Total	Mg	0
			1	1	

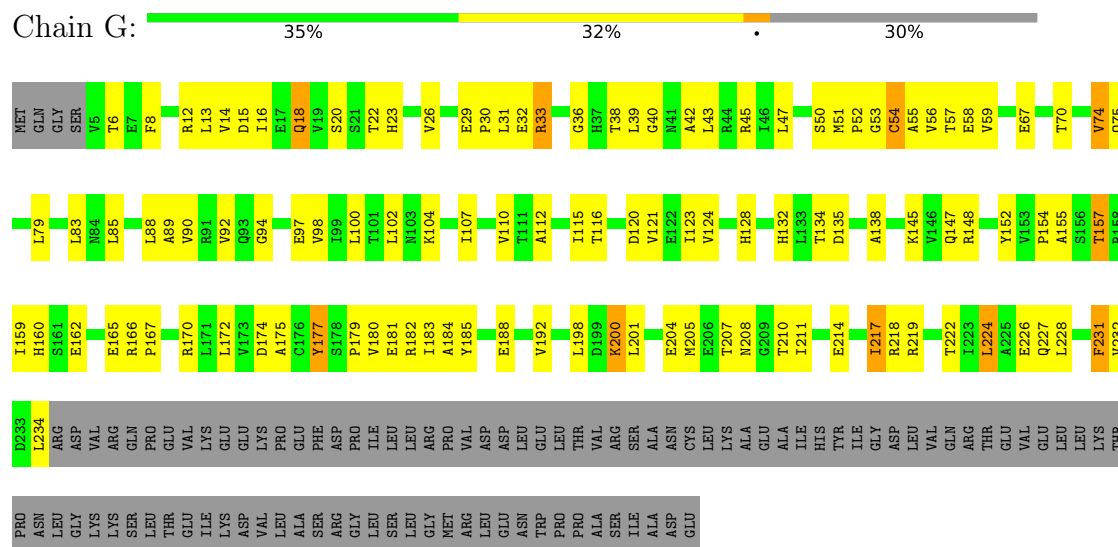
- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
11	J	2	Total 2	Zn 2	0
11	N	1	Total 1	Zn 1	0

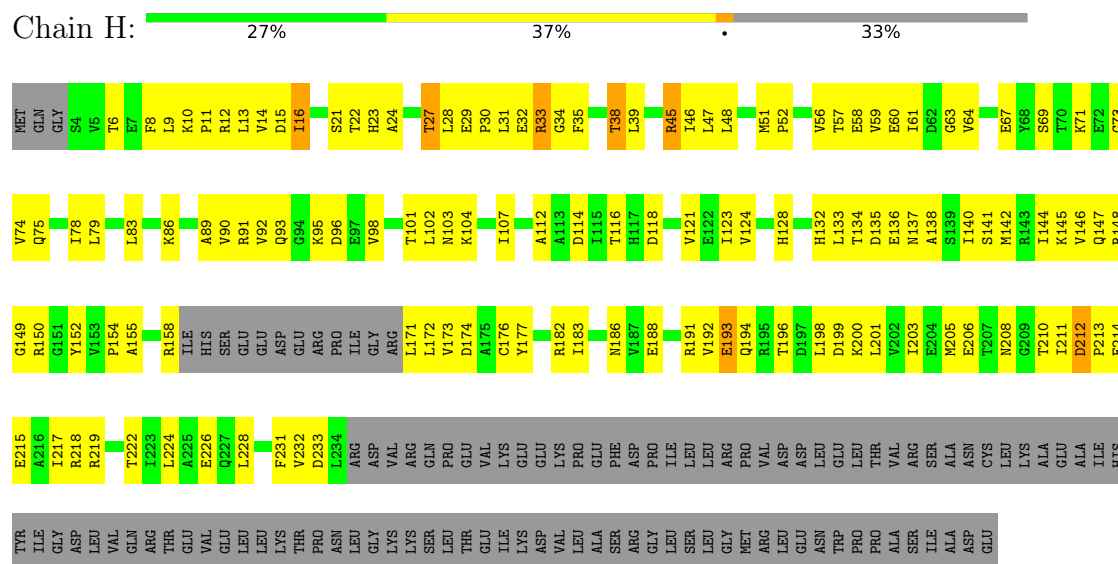
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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 subunit alpha

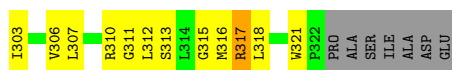
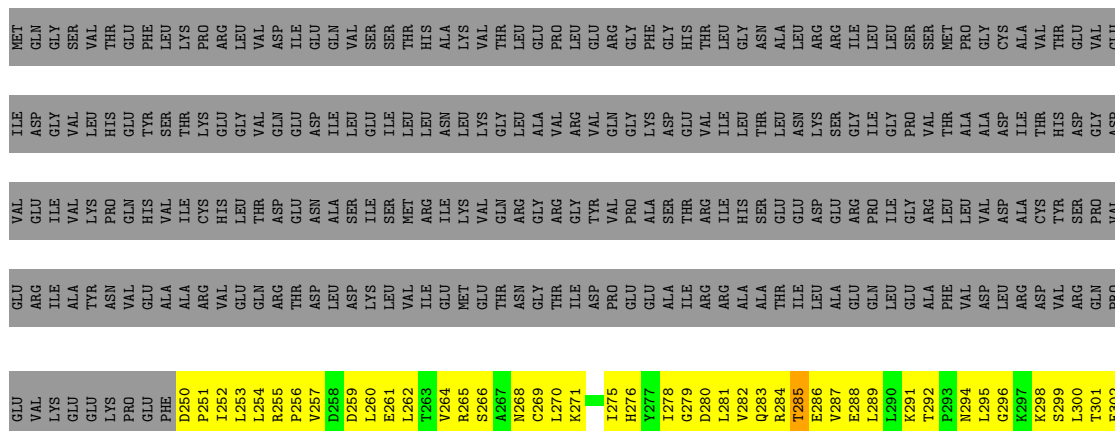


- Molecule 1: DNA-directed RNA polymerase subunit alpha



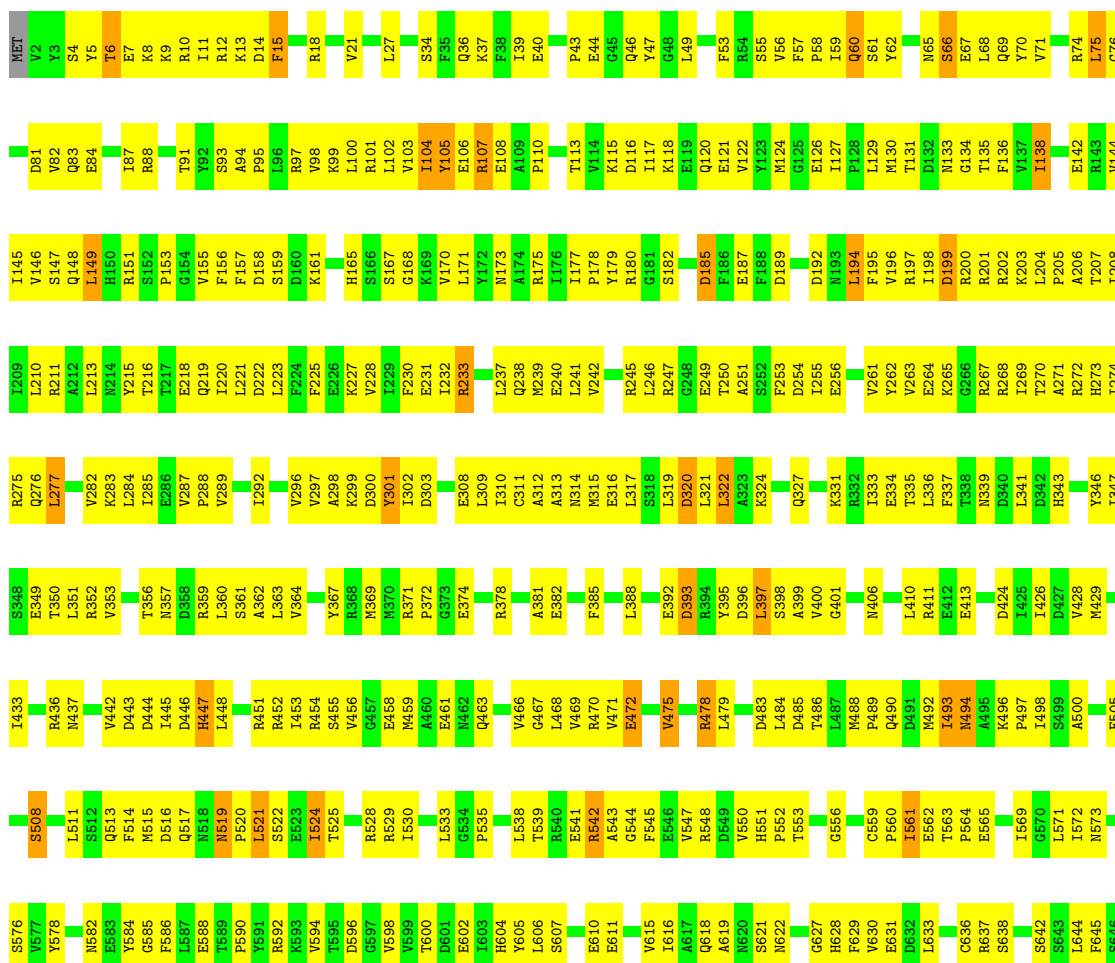
- Molecule 1: DNA-directed RNA polymerase subunit alpha

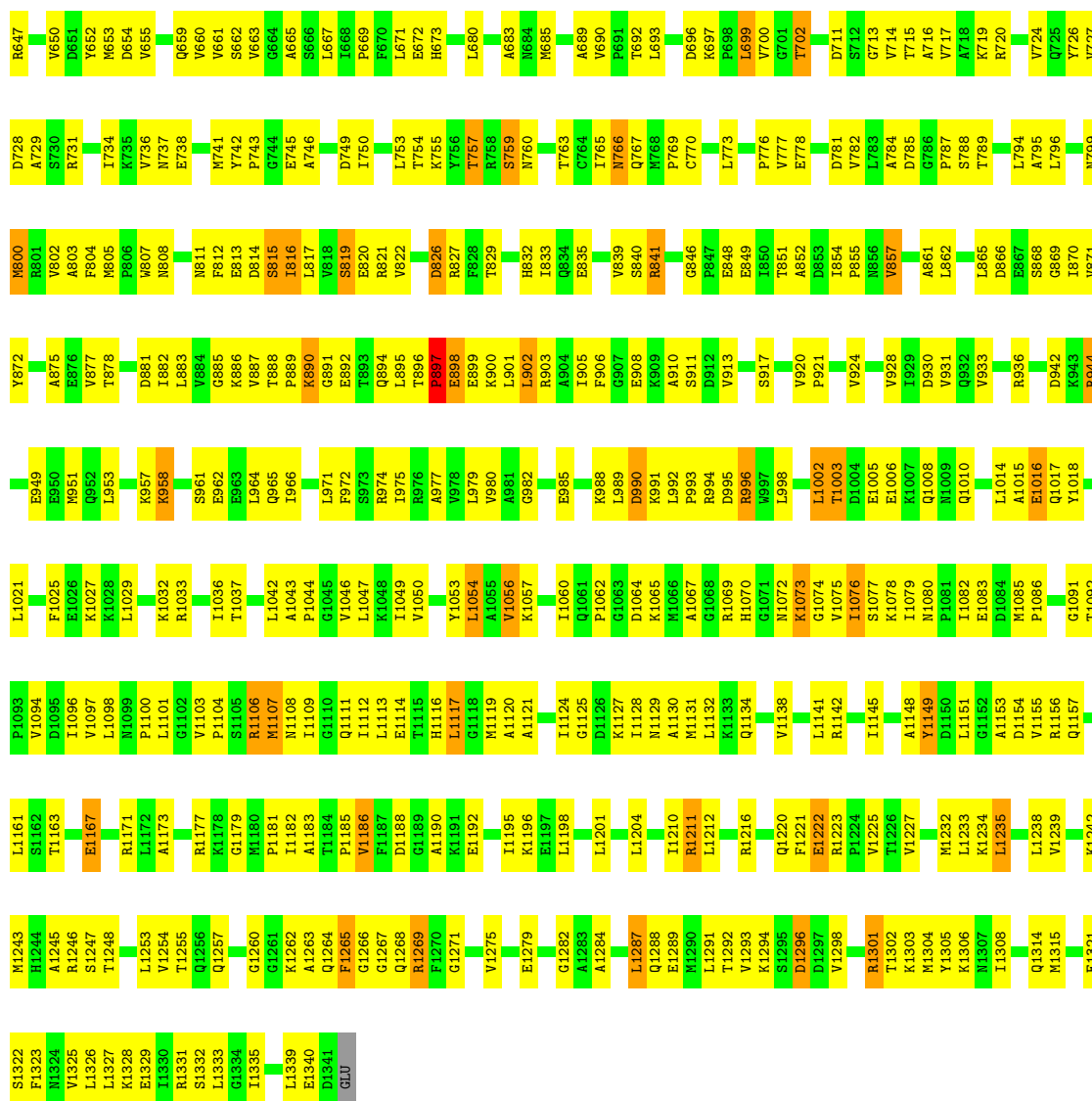
Chain M:  5% 16% . 78%



- Molecule 2: DNA-directed RNA polymerase subunit beta

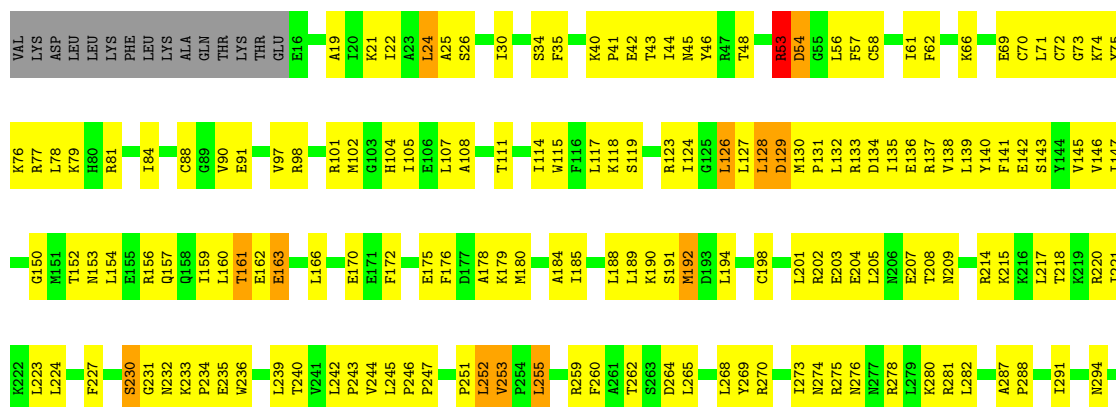
Chain I:  42% 52% 5%





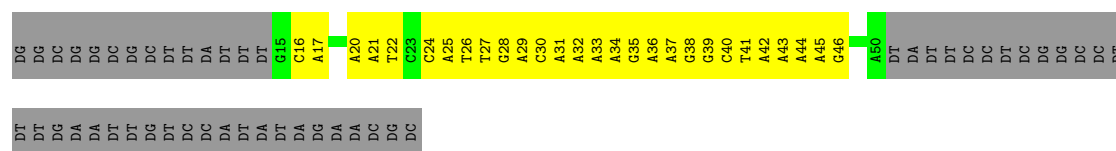
● Molecule 3: DNA-directed RNA polymerase subunit beta'

Chain J: 39% 51% 6%





Chain 0:  9% 33% 58%



- Molecule 8: DNA (85-MER)

Chain P: 39% 59%



4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, C1	Depositor
Number of particles used	98324	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	80	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	GATAN K2 SUMMIT (4k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, 1N7, MG

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	G	0.41	0/1793	0.53	0/2432
1	H	0.37	0/1697	0.54	0/2301
1	M	0.25	0/579	0.47	0/784
2	I	0.43	0/10728	0.52	1/14477 (0.0%)
3	J	0.40	0/10625	0.53	1/14345 (0.0%)
4	K	0.33	0/602	0.49	0/810
5	L	0.29	0/4461	0.48	0/6004
6	N	0.39	0/575	0.49	0/778
7	O	0.61	0/842	0.88	0/1297
8	P	0.58	0/790	1.05	0/1217
All	All	0.41	0/32692	0.55	2/44445 (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
2	I	0	2
3	J	0	1
All	All	0	3

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	J	53	ARG	CB-CA-C	5.29	120.97	110.40
2	I	15	PHE	C-N-CA	-5.09	111.61	122.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	I	519	ASN	Peptide
2	I	897	PRO	Peptide
3	J	53	ARG	Peptide

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	G	1771	0	1799	111	0
1	H	1678	0	1698	131	0
1	M	572	0	602	84	0
2	I	10559	0	10577	767	0
3	J	10466	0	10689	868	0
4	K	600	0	607	48	0
5	L	4407	0	4432	414	0
6	N	565	0	545	35	0
7	O	746	0	401	41	0
8	P	710	0	402	50	0
9	I	27	39	39	4	0
9	J	54	78	75	9	0
9	L	27	39	38	4	0
10	J	1	0	0	0	0
11	J	2	0	0	0	0
11	N	1	0	0	0	0
All	All	32186	156	31904	2388	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 37.

The worst 5 of 2388 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)
9:L:701:1N7:C19	9:L:701:1N7:C3	1.82	1.58
9:I:1401:1N7:C19	9:I:1401:1N7:C3	1.82	1.56
9:J:1505:1N7:C3	9:J:1505:1N7:C19	1.81	1.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
9:J:1504:1N7:C3	9:J:1504:1N7:C19	1.84	1.50
3:J:1179:PRO:HD2	3:J:1184:ASP:HA	1.34	1.08

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 PDB entries followed by that with respect to all EM entries.

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	G	228/329 (69%)	205 (90%)	23 (10%)	0	100	100
1	H	215/329 (65%)	189 (88%)	26 (12%)	0	100	100
1	M	71/329 (22%)	68 (96%)	3 (4%)	0	100	100
2	I	1338/1342 (100%)	1179 (88%)	157 (12%)	2 (0%)	53	85
3	J	1339/1430 (94%)	1197 (89%)	141 (10%)	1 (0%)	53	85
4	K	73/91 (80%)	68 (93%)	5 (7%)	0	100	100
5	L	540/616 (88%)	500 (93%)	40 (7%)	0	100	100
6	N	70/72 (97%)	65 (93%)	5 (7%)	0	100	100
All	All	3874/4538 (85%)	3471 (90%)	400 (10%)	3 (0%)	56	85

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	I	898	GLU
3	J	854	ALA
2	I	897	PRO

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 PDB entries followed by that with respect to all EM entries.

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	G	195/286 (68%)	177 (91%)	18 (9%)	10	37
1	H	184/286 (64%)	171 (93%)	13 (7%)	16	50
1	M	65/286 (23%)	62 (95%)	3 (5%)	29	65
2	I	1153/1157 (100%)	1043 (90%)	110 (10%)	9	36
3	J	1128/1189 (95%)	1036 (92%)	92 (8%)	12	43
4	K	65/75 (87%)	59 (91%)	6 (9%)	10	37
5	L	479/543 (88%)	456 (95%)	23 (5%)	28	64
6	N	60/61 (98%)	53 (88%)	7 (12%)	6	26
All	All	3329/3883 (86%)	3057 (92%)	272 (8%)	17	43

5 of 272 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	I	1106	ARG
3	J	163	GLU
5	L	362	ASN
2	I	1117	LEU
2	I	1287	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
3	J	294	ASN
3	J	489	ASN
5	L	437	GLN
3	J	320	ASN
3	J	465	GLN

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 8 ligands modelled in this entry, 4 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$
9	1N7	I	1401	-	30,30,46	4.86	16 (53%)	47,48,72	2.07	10 (21%)
9	1N7	J	1504	-	30,30,46	4.99	16 (53%)	47,48,72	2.37	17 (36%)
9	1N7	J	1505	-	30,30,46	4.73	14 (46%)	47,48,72	2.10	14 (29%)
9	1N7	L	701	-	30,30,46	4.79	15 (50%)	47,48,72	2.46	20 (42%)

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
9	1N7	I	1401	-	-	5/7/72/92	0/4/4/4
9	1N7	J	1504	-	-	7/7/72/92	0/4/4/4
9	1N7	J	1505	-	-	0/7/72/92	0/4/4/4
9	1N7	L	701	-	-	5/7/72/92	0/4/4/4

The worst 5 of 61 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	J	1504	1N7	C3-C19	18.15	1.84	1.53

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	I	1401	1N7	C3-C19	17.43	1.82	1.53
9	L	701	1N7	C3-C19	16.95	1.82	1.53
9	J	1505	1N7	C3-C19	16.87	1.81	1.53
9	J	1504	1N7	C3-C4	11.42	1.73	1.53

The worst 5 of 61 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	L	701	1N7	C9-C5-C4	-7.70	110.59	117.68
9	J	1504	1N7	C9-C5-C4	-7.57	110.71	117.68
9	I	1401	1N7	C9-C5-C4	-6.62	111.58	117.68
9	L	701	1N7	C3-C19-C18	-6.36	101.58	110.86
9	J	1505	1N7	C19-C3-C4	-4.92	107.77	114.30

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
9	J	1504	1N7	C21-C20-C9-C5
9	J	1504	1N7	C21-C20-C9-C8
9	I	1401	1N7	C21-C20-C9-C5
9	I	1401	1N7	C21-C20-C9-C8
9	J	1504	1N7	C22-C20-C9-C5

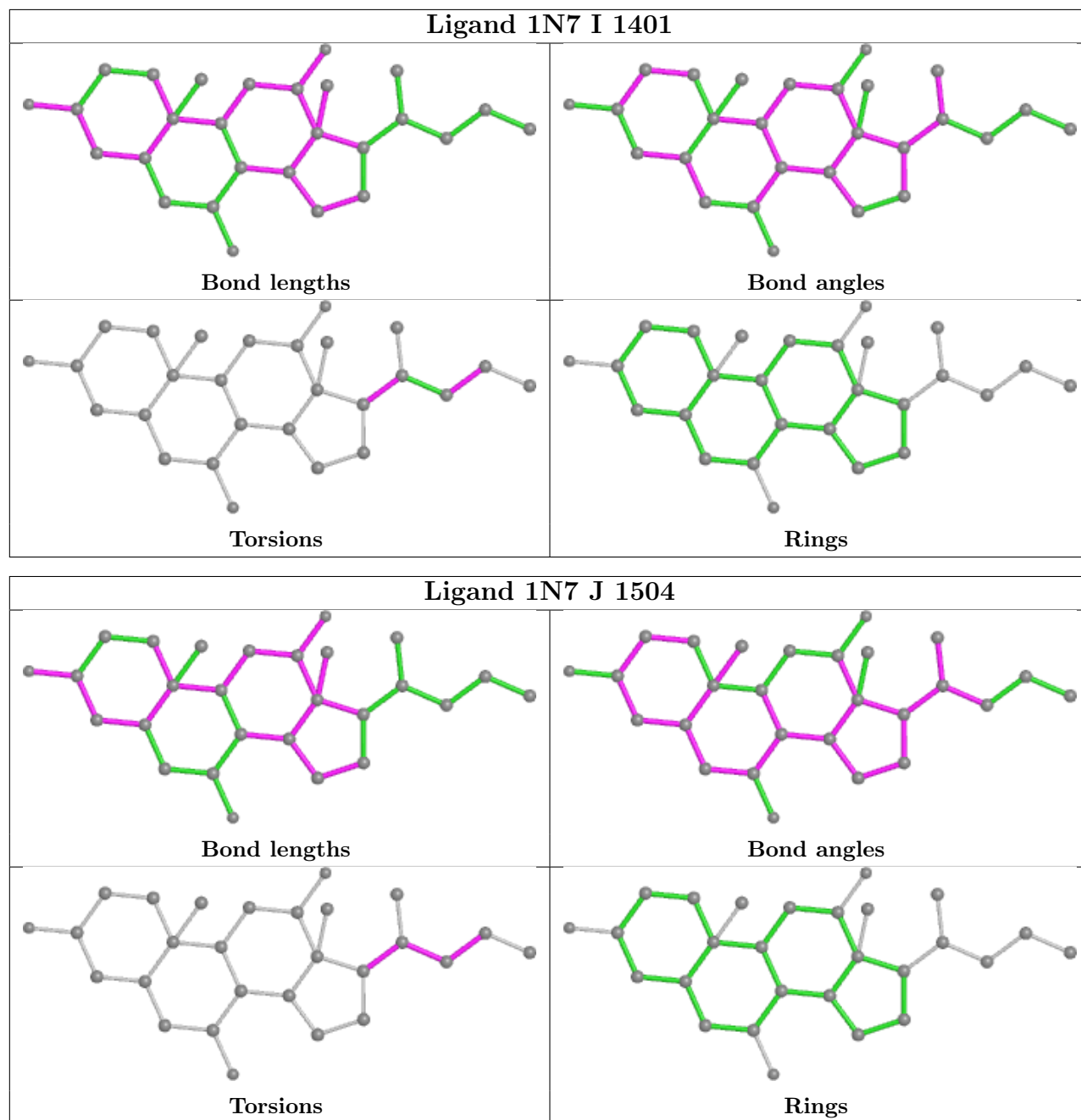
There are no ring outliers.

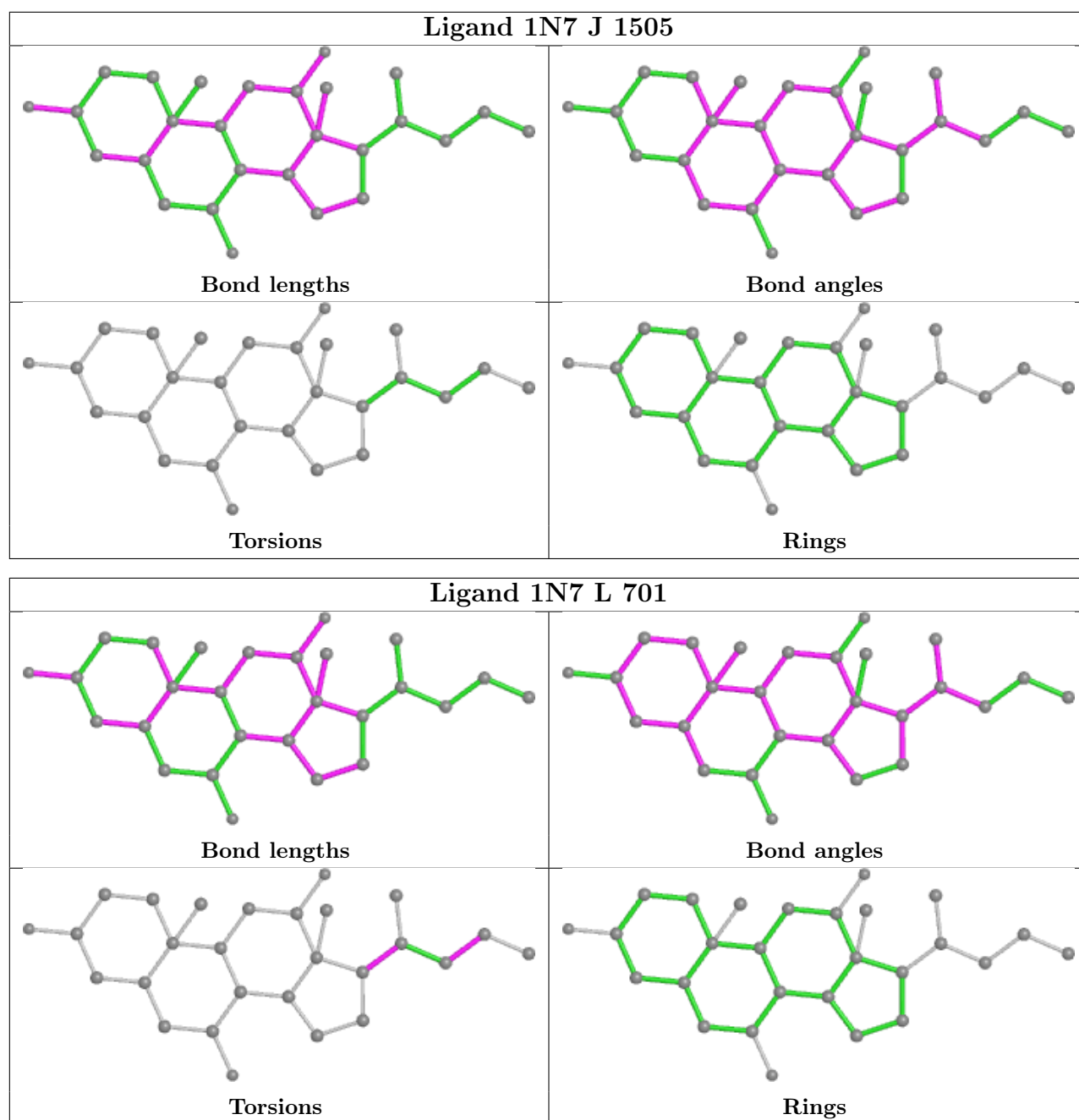
4 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	I	1401	1N7	4	0
9	J	1504	1N7	4	0
9	J	1505	1N7	5	0
9	L	701	1N7	4	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and

any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.





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.