wwPDB X-ray Structure Validation Summary Report (i)

Oct 5, 2023 – 07:26 PM EDT

PDB ID	:	6UPZ
Title	:	RNA polymerase II elongation complex with 5-guanidinohydantoin lesion in state 3
Authors	:	Oh, J.; Wang, D.
Deposited on	:	2019-10-18
Resolution	:	3.10 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

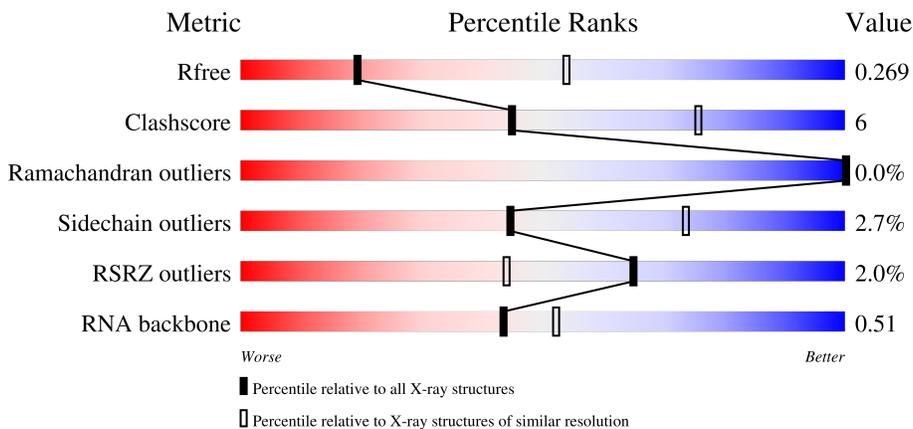|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 3.10 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	Percentile Ranks	Value
Rfree		0.269
Clashscore		6
Ramachandran outliers		0.0%
Sidechain outliers		2.7%
RSRZ outliers		2.0%
RNA backbone		0.51

Worse — *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)
RNA backbone	3102	1116 (3.40-2.80)

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% 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	R	10	30% 70%
2	T	29	10% / 69% 17% · 10%
3	N	18	11% / 67% 17% 17%
4	A	1733	2% / 64% 15% · 20%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	B	1224	% 75% 16% • 8%
6	C	318	69% 14% 16%
7	E	215	8% 80% 19% •
8	F	155	46% 8% • 45%
9	H	146	4% 68% 22% • 9%
10	I	122	85% 11% •
11	J	70	71% 20% • 7%
12	K	120	84% 11% 5%
13	L	70	3% 53% 9% 39%

2 Entry composition ⓘ

There are 16 unique types of molecules in this entry. The entry contains 29075 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 RNA chain called RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	10	Total	C	N	O	P	0	0	0
			221	98	45	68	10			

- Molecule 2 is a DNA chain called Template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	26	Total	C	N	O	P	0	0	0
			520	249	80	165	26			

- Molecule 3 is a DNA chain called Non-template strand DNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	15	Total	C	N	O	P	0	0	0
			317	148	71	83	15			

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1382	Total	C	N	O	S	0	0	0
			10820	6824	1895	2041	60			

- Molecule 5 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1123	Total	C	N	O	S	0	0	0
			8859	5607	1552	1647	53			

- Molecule 6 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	C	267	Total	C	N	O	S	0	0	0
			2101	1320	349	419	13			

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	E	212	Total 1731	C 1100	N 305	O 315	S 11	0	0	0

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	F	86	Total 684	C 437	N 115	O 129	S 3	0	0	0

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	H	133	Total 1064	C 670	N 179	O 211	S 4	0	0	0

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	I	118	Total 952	C 585	N 173	O 184	S 10	0	0	0

- Molecule 11 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

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

- Molecule 12 is a protein called DNA-directed RNA polymerase II subunit RPB11.

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

- Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	L	43	Total 337	C 208	N 66	O 59	S 4	0	0	0

- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 16 is PYROPHOSPHATE (three-letter code: PPV) (formula: $H_4O_7P_2$) (labeled as "Ligand of Interest" by depositor).

PPV
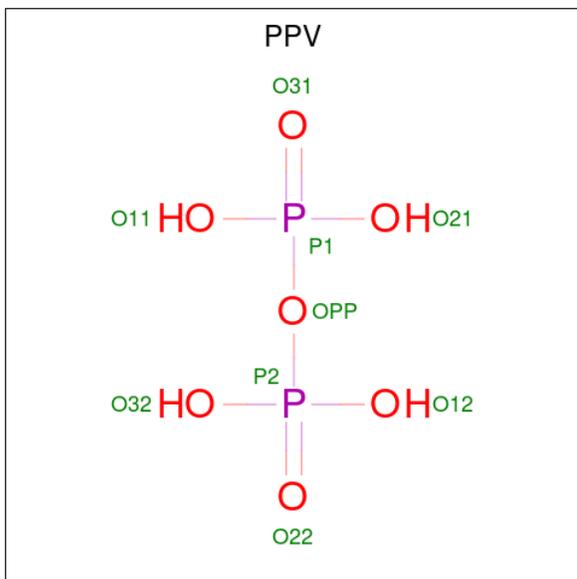
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
16	B	1	Total 9	O 7	P 2	0	0

3 Residue-property plots ⓘ

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: RNA
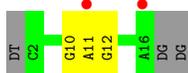
Chain R:

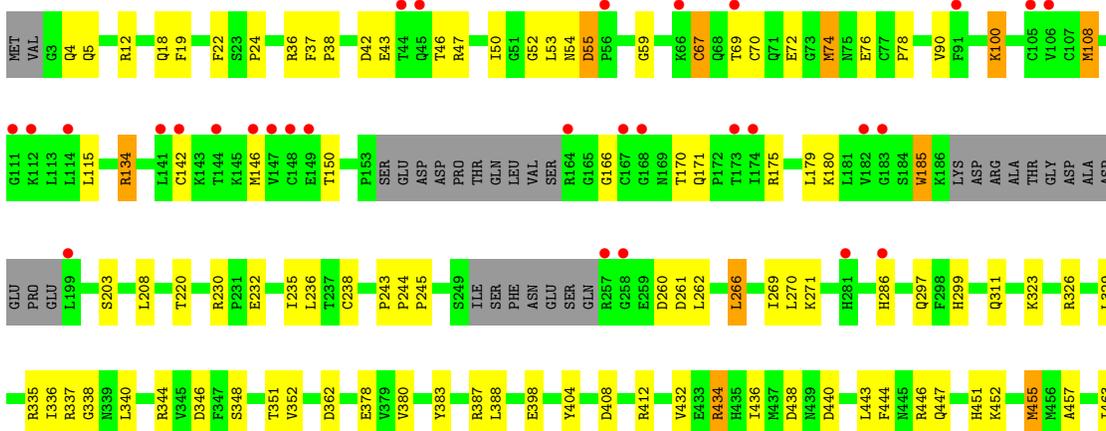• Molecule 2: Template strand DNA

Chain T:
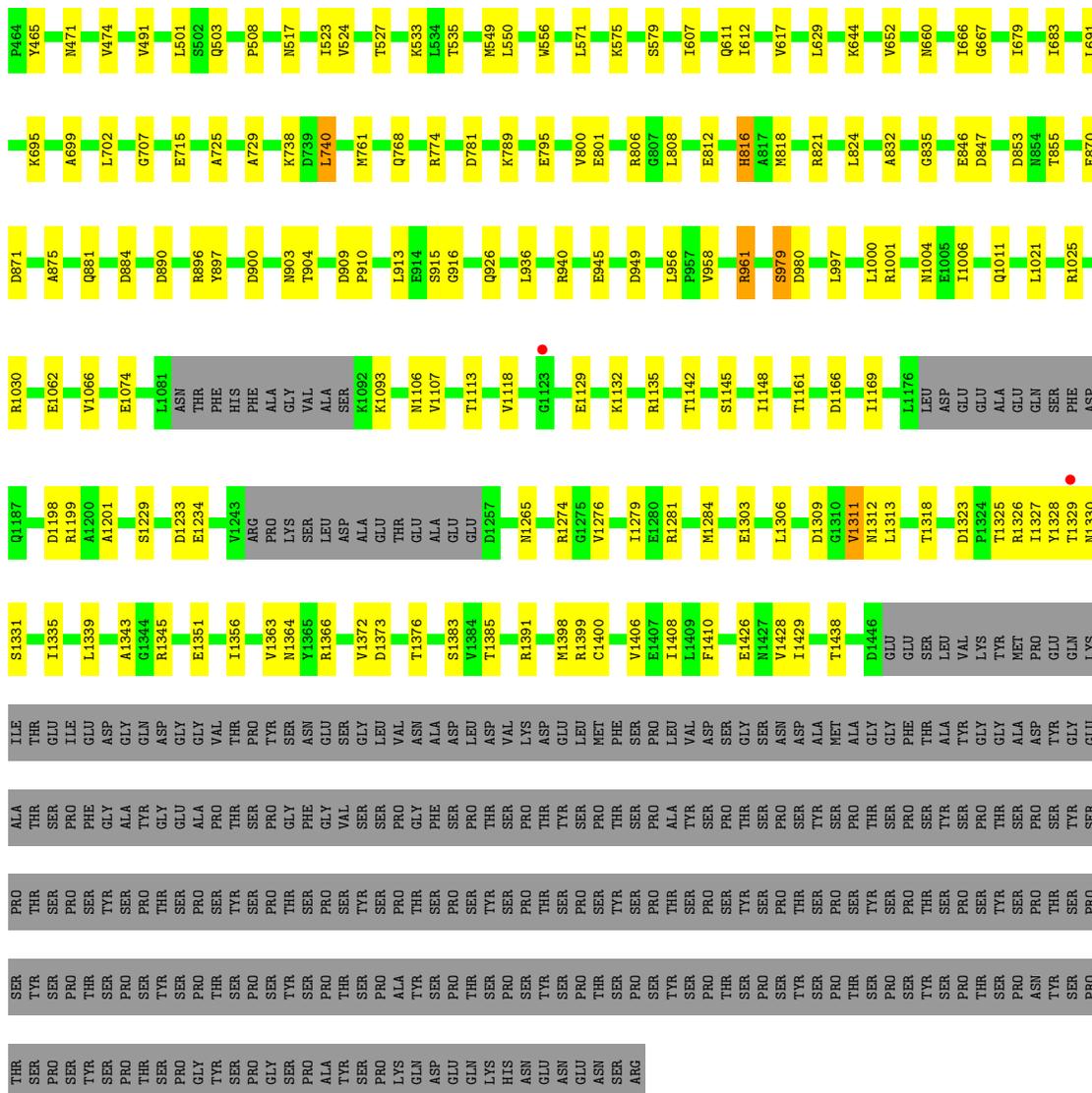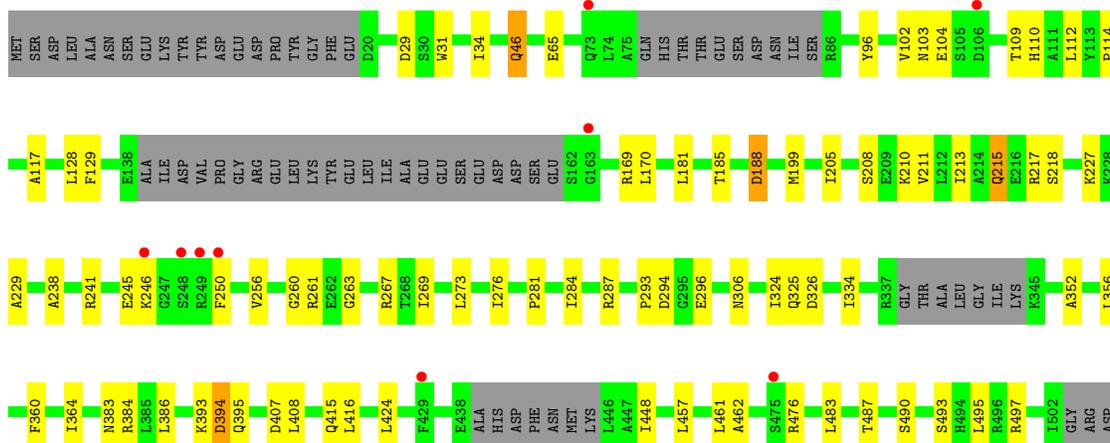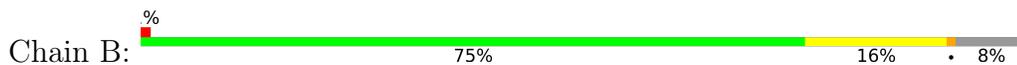
• Molecule 3: Non-template strand DNA

Chain N:
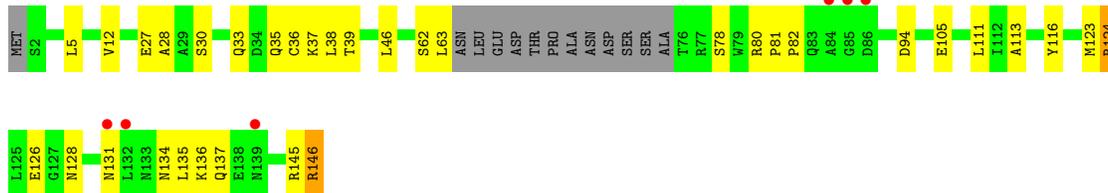
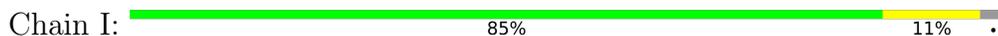• Molecule 4: DNA-directed RNA polymerase II subunit RPB1

Chain A:

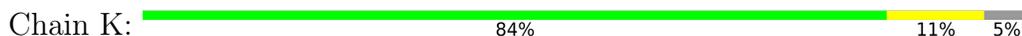• Molecule 5: DNA-directed RNA polymerase II subunit RPB2

Chain B:

GLY LYS LEU A509 L514 H518 V522 T549 A577 V585 V589 N610 P611 E612 V613 S614 M615 I626 F627 T628 D629 R632 I639 D643 GLU SER LEU G647 V653 R654 K655 I658 D668 ILE GLU GLY GLY PHE GLU ASP V676 L689 V690 E691

M705 E714 ALA ASN GLU GLU ASN ASP L721 D722 P725 R728 V731 S732 H733 H744 P745 S746 D760 Q763 S764 P765 R766 M778 R788 M789 D790 A793 N794 I795 L796 P799 T805 T806 R807 A808 R815 Y830 N834 Q835 E836 D837 I840

Q843 D847 R848 F856 R857 Q862 E863 K864 K865 Y866 S869 R884 Y890 D895 A900 P901 R904 I912 S919 PRO ASP GLU GLU LEU LEU GLY GLN ARG THR A930 K934 R935 D936 T939 V949 L953 N957 L961 G977 D978 K979

I990 G991 I992 T993 Y994 R995 R996 E997 D998 M999 P1000 F1001 T1002 A1003 E1004 I1011 I1012 N1013 R1020 M1021 T1022 V1023 L1026 I1027 S1056 R1060 S1066 M1072 M1082 L1095 V1099 D1100 D1101 K1102 I1103 H1104 A1105 R1106 A1107 R1108 M1111 R1116 R1122 G1126 G1127

L1128 R1129 R1135 D1136 H1141 L1147 L1150 L1151 R1159 K1174 Q1179 Q1194 K1183 Q1193 I1194 H1195 Q1205 S1221 ARG ASP PHE

- Molecule 6: DNA-directed RNA polymerase II subunit RPB3

Chain C: 69% 14% 16%

MET S2 E3 P6 I10 S14 F20 I21 L22 V25 R34 R35 V36 M37 E40 I41 P42 T43 L44 A45 I46 E52 T53 N54 D60 E61 F62 R66 I77 D93 Y114 R127 N140 G141 V142 L143 I144 K154 A159 K160 K161 K165

E166 H167 A174 E177 F178 E179 W192 N206 D217 N231 V244 K254 V255 A256 L259 D268 LYS VAL ASN PHE ALA SER GLY ASP ASN ASN THR ALA SER ASN MET LEU GLY SER ASN GLU ASP VAL MET MET THR GLY ALA GLU ASP GLN ASP PRO TYR SER ASN

ALA SER GLN MET GLY ASN THR GLY SER GLY TYR ASP ASN ALA TRP

- Molecule 7: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E: 8% 80% 19%

MET ASP GLN E4 R7 S10 R11 K20 K24 Q32 E33 E34 L39 K43 Y46 C47 R55 K56 M57 M58 S59 P64 G76 E81 F82 C83 V88 K91 T92 M93 K94 I100 F110 I116 T117 P118 S119 A120 M121 K122 L123 V124 P125

S126 I127 P128 A130 F135 N136 E137 L140 V141 V142 H146 H147 V150 E163 R167 K197 I198 I199 R200 K201 S205 M215

- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F: 46% 8% 45%

MET SER ASP TYR GLU GLU ALA PHE ASN ASP GLY ASN GLU ASN PHE GLU ASP PHE ASP VAL GLU HIS PHE SER ASP GLU GLU THR TYR GLU GLU LYS PRO GLN PHE LYS ASP GLY GLU THR ASP ALA ASN GLY LYS THR ILE VAL THR GLY GLY ASN GLY PRO GLU ASP PHE GLN GLN

HIS GLU GLN ILE ARG ARG LYS THR L69 K76 D77 Q78 R79 M85 Y88 E89 R90 A91 R92 I93 R97 L111 K128 V148 E149 E150 D154 LEU

- Molecule 9: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:

4%

68%   22%   •   9%

MET S2 L5 V12 E27 A28 A29 S30 Q33 D34 Q35 C36 K37 L38 T39 L46 S62 L63 ASN LEU GLU ASP THR PRO ALA ASN ASP SER ALA T76 R77 S78 W79 R80 P81 P82 Q83 A84 G85 D86 D94 E105 L111 I112 A113 Y116 M123 R124

L125 E126 G127 N128 N131 L132 N133 N134 L135 K136 Q137 E138 N139 R145 R146

• Molecule 10: DNA-directed RNA polymerase II subunit RPB9

Chain I:

85%   11%   •

MET T2 T3 F4 R5 F6 C7 N12 E36 I49 T50 N51 I52 R73 N83 R91 T111 S112 D113 Q114 T119 GLN PHE SER

• Molecule 11: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J:

71%   20%   •   7%

M1 I2 V3 P4 V5 F8 S9 V13 V14 K17 L36 S37 R38 L39 Y44 R47 R48 M49 I50 L51 P65 LEU GLU LYS ARG ASP

• Molecule 12: DNA-directed RNA polymerase II subunit RPB11

Chain K:

84%   11%   5%

M1 R6 F10 L11 D24 T25 K26 V32 T41 E49 R54 R74 E79 G80 Y81 I94 L114 ALA ALA ASP ASP ALA PHE

• Molecule 13: DNA-directed RNA polymerases I, II, and III subunit RPABC4

Chain L:

3%

53%   9%   39%

MET SER ARG GLU GLY PHE GLN ILE PRO THR ASN LEU ASP ALA ALA ALA ALA GLY THR SER GLN ALA ARG THR ALA THR LEU K28 C31 A32 S41 R42 T43 K49 D50 T61 Q66 R70

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	167.66Å 222.90Å 193.49Å 90.00° 100.85° 90.00°	Depositor
Resolution (Å)	49.04 – 3.10 49.04 – 3.10	Depositor EDS
% Data completeness (in resolution range)	100.0 (49.04-3.10) 100.0 (49.04-3.10)	Depositor EDS
R_{merge}	0.36	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.51 (at 3.12Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.224 , 0.269 0.224 , 0.269	Depositor DCC
R_{free} test set	1849 reflections (1.47%)	wwPDB-VP
Wilson B-factor (Å^2)	64.3	Xtriage
Anisotropy	0.471	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 49.3	EDS
L-test for twinning[2]	$< |L| > = 0.46, < L^2 > = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	29075	wwPDB-VP
Average B, all atoms (Å^2)	85.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.74% 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: PPV, MG, G35, ZN

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	R	0.72	1/248 (0.4%)	0.79	0/384
2	T	0.63	0/551	1.06	0/843
3	N	0.65	0/359	0.84	0/553
4	A	0.28	0/11012	0.49	2/14895 (0.0%)
5	B	0.29	0/9030	0.48	0/12186
6	C	0.30	0/2139	0.49	0/2899
7	E	0.27	0/1767	0.48	0/2378
8	F	0.27	0/696	0.47	0/943
9	H	0.27	0/1082	0.51	0/1466
10	I	0.30	0/970	0.50	0/1308
11	J	0.29	0/541	0.48	0/727
12	K	0.28	0/937	0.48	0/1265
13	L	0.34	0/339	0.52	0/450
All	All	0.31	1/29671 (0.0%)	0.52	2/40297 (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
4	A	0	5

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	R	1	A	OP3-P	-10.56	1.48	1.61

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	398	GLU	C-N-CA	6.14	137.05	121.70
4	A	266	LEU	CA-CB-CG	5.05	126.92	115.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	1311	VAL	Peptide
4	A	311	GLN	Peptide
4	A	524	VAL	Peptide
4	A	53	LEU	Peptide
4	A	55	ASP	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	R	221	0	110	3	0
2	T	520	0	299	8	0
3	N	317	0	166	2	0
4	A	10820	0	10862	156	0
5	B	8859	0	8816	138	0
6	C	2101	0	2056	29	0
7	E	1731	0	1758	22	0
8	F	684	0	692	13	0
9	H	1064	0	1029	17	0
10	I	952	0	899	9	0
11	J	532	0	542	10	0
12	K	919	0	929	15	0
13	L	337	0	352	5	0
14	A	2	0	0	0	0
14	B	1	0	0	0	0
14	C	1	0	0	0	0
14	I	2	0	0	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0
15	A	1	0	0	0	0
16	B	9	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	29075	0	28510	370	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 6.

The worst 5 of 370 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)
5:B:213:ILE:O	5:B:215:GLN:NE2	1.90	1.02
2:T:19:G35:H1'	4:A:835:GLY:HA3	1.62	0.81
5:B:999:MET:HE2	5:B:1000:PRO:HD2	1.63	0.81
4:A:262:LEU:O	4:A:266:LEU:HD23	1.86	0.76
6:C:66:ARG:NH2	11:J:3:VAL:O	2.21	0.74

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	A	1368/1733 (79%)	1323 (97%)	44 (3%)	1 (0%)	51	83
5	B	1103/1224 (90%)	1070 (97%)	33 (3%)	0	100	100
6	C	265/318 (83%)	259 (98%)	6 (2%)	0	100	100
7	E	210/215 (98%)	202 (96%)	8 (4%)	0	100	100
8	F	84/155 (54%)	82 (98%)	2 (2%)	0	100	100
9	H	129/146 (88%)	121 (94%)	8 (6%)	0	100	100
10	I	116/122 (95%)	112 (97%)	4 (3%)	0	100	100
11	J	63/70 (90%)	63 (100%)	0	0	100	100
12	K	112/120 (93%)	109 (97%)	3 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
13	L	41/70 (59%)	39 (95%)	2 (5%)	0	100	100
All	All	3491/4173 (84%)	3380 (97%)	110 (3%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	55	ASP

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	
4	A	1193/1520 (78%)	1152 (97%)	41 (3%)	37	69
5	B	955/1061 (90%)	935 (98%)	20 (2%)	53	79
6	C	235/274 (86%)	232 (99%)	3 (1%)	69	87
7	E	193/197 (98%)	188 (97%)	5 (3%)	46	74
8	F	73/137 (53%)	71 (97%)	2 (3%)	44	74
9	H	116/128 (91%)	109 (94%)	7 (6%)	19	49
10	I	110/116 (95%)	108 (98%)	2 (2%)	59	82
11	J	60/65 (92%)	59 (98%)	1 (2%)	60	83
12	K	99/102 (97%)	98 (99%)	1 (1%)	76	90
13	L	37/57 (65%)	37 (100%)	0	100	100
All	All	3071/3657 (84%)	2989 (97%)	82 (3%)	44	74

5 of 82 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	1060	ARG
9	H	33	GLN
5	B	1082	MET
7	E	11	ARG
9	H	131	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	1015	HIS
5	B	1074	ASN
5	B	1195	HIS
4	A	515	GLN
4	A	390	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	9/10 (90%)	1 (11%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	10	A

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	#$|Z| > 2$
2	G35	T	19	2	18,23,24	4.70	14 (77%)	20,33,36	2.11	4 (20%)

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
2	G35	T	19	2	-	4/10/41/42	0/2/2/2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	19	G35	C2-N3	9.37	1.49	1.33
2	T	19	G35	O4'-C4'	7.75	1.62	1.45
2	T	19	G35	C3'-C4'	-7.19	1.33	1.53
2	T	19	G35	C5-N7	6.76	1.46	1.37
2	T	19	G35	C8-N9	6.26	1.46	1.37

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	T	19	G35	C2'-C1'-N9	-5.82	107.73	115.59
2	T	19	G35	C5-C4-N9	4.33	108.03	102.28
2	T	19	G35	O4'-C1'-N9	3.80	113.15	108.65
2	T	19	G35	O4'-C4'-C3'	-2.85	99.03	105.67

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	T	19	G35	O4'-C1'-N9-C4
2	T	19	G35	C3'-C4'-C5'-O5'
2	T	19	G35	O4'-C4'-C5'-O5'
2	T	19	G35	C4'-C5'-O5'-P

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	T	19	G35	3	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 10 ligands modelled in this entry, 9 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	$\#|Z| > 2$
16	PPV	B	1301	-	6,8,8	0.74	0	13,13,13	1.09	1 (7%)

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	PPV	B	1301	-	-	0/6/6/6	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
16	B	1301	PPV	P2-OPP-P1	-2.44	124.45	132.83

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 then 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.

Ligand PPV B 1301

Bond lengths

Bond angles

Torsions

Rings
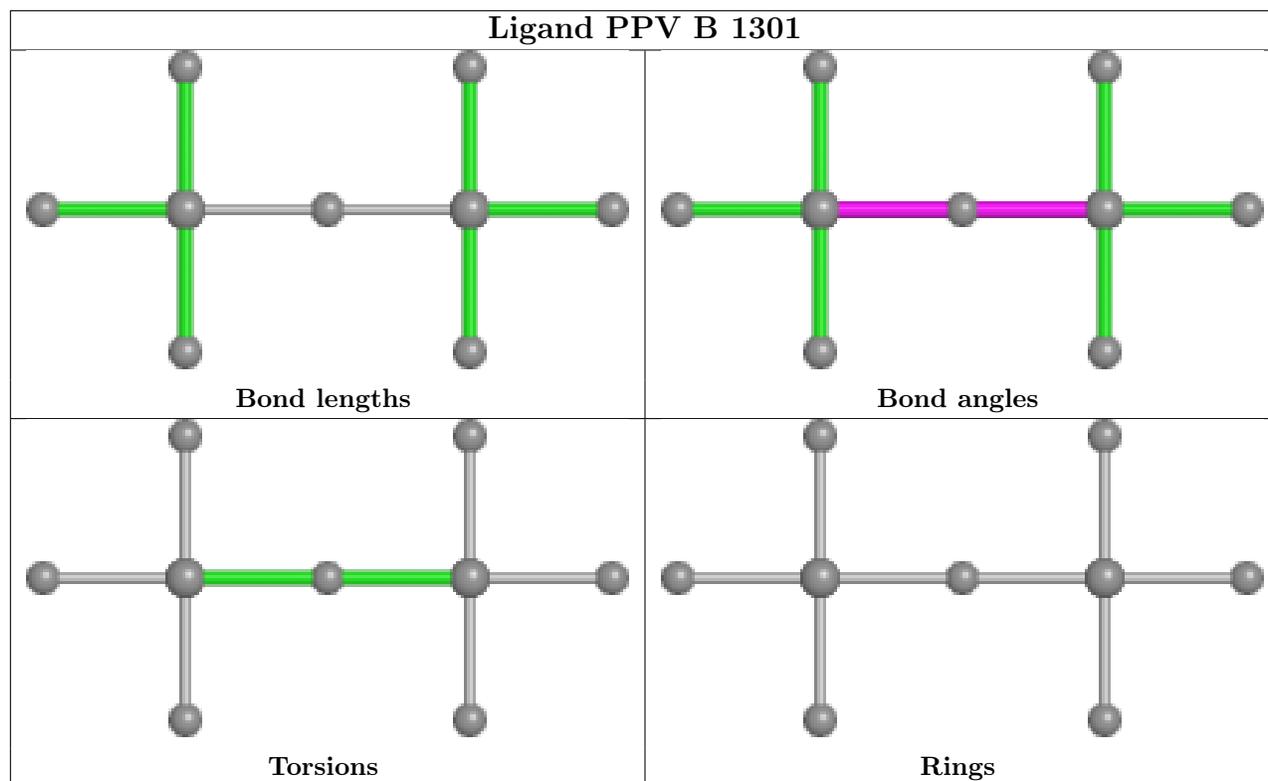
5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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, 95th 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(Å²)	Q<0.9
1	R	10/10 (100%)	-0.39	0	100	100	94, 113, 153, 166	0
2	T	25/29 (86%)	0.42	3 (12%)	4	2	100, 180, 235, 238	0
3	N	15/18 (83%)	0.62	2 (13%)	3	1	142, 179, 248, 270	0
4	A	1382/1733 (79%)	-0.12	32 (2%)	60	39	39, 79, 157, 254	0
5	B	1123/1224 (91%)	-0.22	11 (0%)	82	67	32, 64, 125, 196	0
6	C	267/318 (83%)	-0.44	0	100	100	32, 64, 102, 141	0
7	E	212/215 (98%)	0.17	17 (8%)	12	5	55, 113, 190, 220	0
8	F	86/155 (55%)	-0.37	0	100	100	47, 78, 121, 167	0
9	H	133/146 (91%)	0.19	6 (4%)	33	16	68, 99, 149, 201	0
10	I	118/122 (96%)	-0.32	0	100	100	49, 82, 116, 170	0
11	J	65/70 (92%)	-0.45	0	100	100	37, 54, 89, 132	0
12	K	114/120 (95%)	-0.36	0	100	100	31, 67, 101, 140	0
13	L	43/70 (61%)	0.13	2 (4%)	31	15	45, 111, 159, 189	0
All	All	3593/4230 (84%)	-0.16	73 (2%)	65	44	31, 75, 156, 270	0

The worst 5 of 73 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	A	69	THR	8.8
4	A	44	THR	6.9
5	B	250	PHE	6.1
7	E	126	SER	5.7
7	E	91	LYS	5.6

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

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, 95th 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^2)	Q<0.9
2	G35	T	19	22/23	0.70	0.39	146,186,196,200	0

6.3   Carbohydrates ⓘ
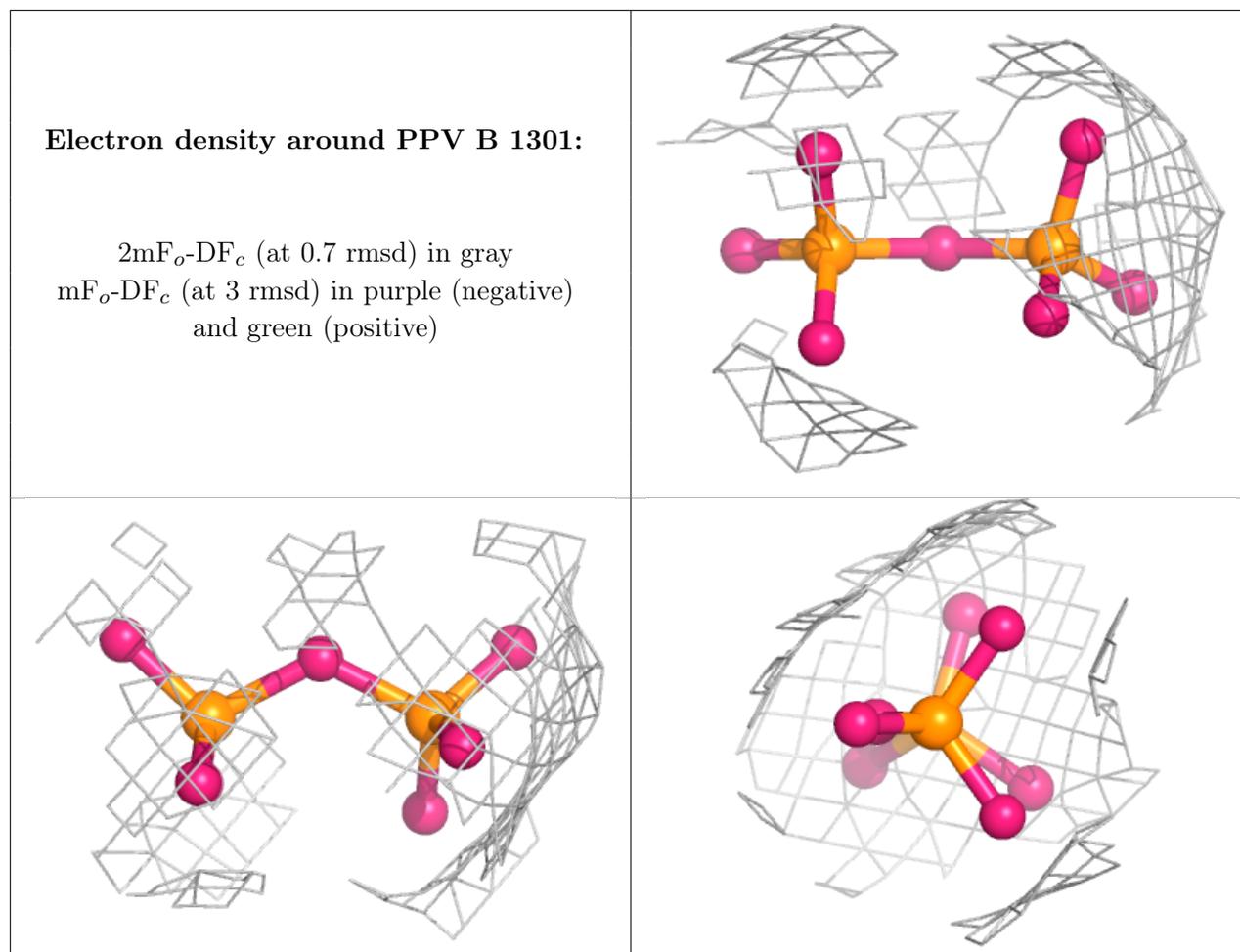
There are no monosaccharides in this entry.

6.4   Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th 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^2)	Q<0.9
14	ZN	C	401	1/1	0.87	0.34	243,243,243,243	0
16	PPV	B	1301	9/9	0.89	0.26	130,140,155,155	0
14	ZN	A	1801	1/1	0.91	0.07	194,194,194,194	0
14	ZN	A	1802	1/1	0.96	0.08	105,105,105,105	0
14	ZN	L	101	1/1	0.96	0.05	137,137,137,137	0
14	ZN	B	1302	1/1	0.96	0.13	124,124,124,124	0
14	ZN	J	101	1/1	0.97	0.31	111,111,111,111	0
15	MG	A	1803	1/1	0.98	0.09	47,47,47,47	0
14	ZN	I	201	1/1	0.99	0.12	56,56,56,56	0
14	ZN	I	202	1/1	0.99	0.13	55,55,55,55	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around PPV B 1301:

$2mF_o\text{-}DF_c$ (at 0.7 rmsd) in gray
$mF_o\text{-}DF_c$ (at 3 rmsd) in purple (negative)
and green (positive)

6.5   Other polymers ⓘ

There are no such residues in this entry.