



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 02:52 PM UTC

PDB ID : 5VSW / pdb_00005vsw
Title : X-ray crystal structure of Escherichia coli RNA polymerase and DksA/ppGpp complex
Authors : Murakami, K.S.; Molodtsov, V.
Deposited on : 2017-05-12
Resolution : 4.29 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 ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Buster-report	:	wwPDB partial adaption of 1.1.7 (2018)
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

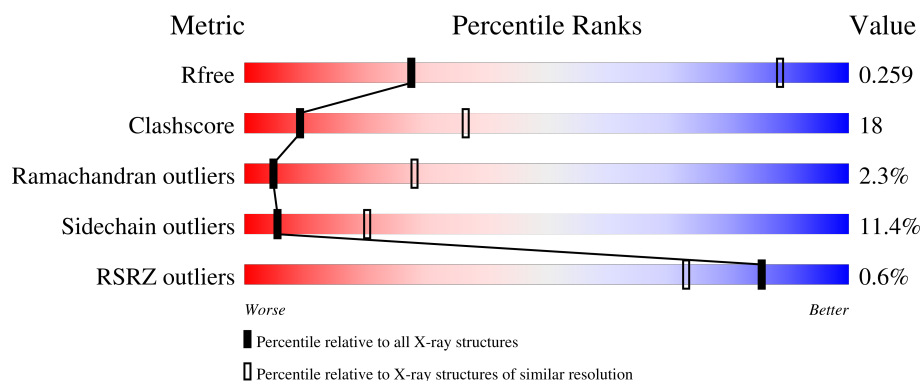
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.29 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}	180053	1052 (4.70-3.90)
Clashscore	190562	1097 (4.70-3.90)
Ramachandran outliers	187476	1001 (4.70-3.90)
Sidechain outliers	187428	1007 (4.72-3.88)
RSRZ outliers	180081	1049 (4.70-3.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	329	50% 36% 11% ••
1	B	329	43% 39% 7% • 10%
1	G	329	% 41% 24% • 31%
1	H	329	39% 24% • 34%
2	C	1342	% 58% 36% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	I	1342	% 60% 35% 5%
3	D	1407	% 47% 32% 17%
3	J	1407	% 46% 32% 18%
4	E	91	% 60% 33% 5% 2%
4	K	91	% 52% 26% 8% 13%
5	F	613	% 46% 27% 24%
5	L	613	% 46% 27% 23%
6	M	151	% 45% 44% 7%

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 57643 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 subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	319	Total	C	N	O	S	0	0	0
			2490	1557	439	486	8			
1	B	297	Total	C	N	O	S	0	0	0
			2297	1439	403	447	8			
1	G	227	Total	C	N	O	S	0	0	0
			1755	1093	311	345	6			
1	H	216	Total	C	N	O	S	0	0	0
			1662	1038	292	326	6			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1340	Total	C	N	O	S	0	0	0
			10570	6631	1841	2055	43			
2	I	1340	Total	C	N	O	S	0	0	0
			10566	6629	1840	2054	43			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1173	Total	C	N	O	S	0	0	0
			9095	5718	1628	1703	46			
3	J	1155	Total	C	N	O	S	0	0	0
			9001	5659	1612	1684	46			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	89	Total	C	N	O	S	0	0	0
			691	421	129	140	1			
4	K	79	Total	C	N	O	S	0	0	0
			627	382	118	126	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	468	Total	C	N	O	S	0	0	0
			3813	2389	678	723	23			
5	L	469	Total	C	N	O	S	0	0	0
			3821	2393	679	726	23			

- Molecule 6 is a protein called RNA polymerase-binding transcription factor DksA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	M	140	Total	C	N	O	S	0	0	0
			1140	703	206	224	7			

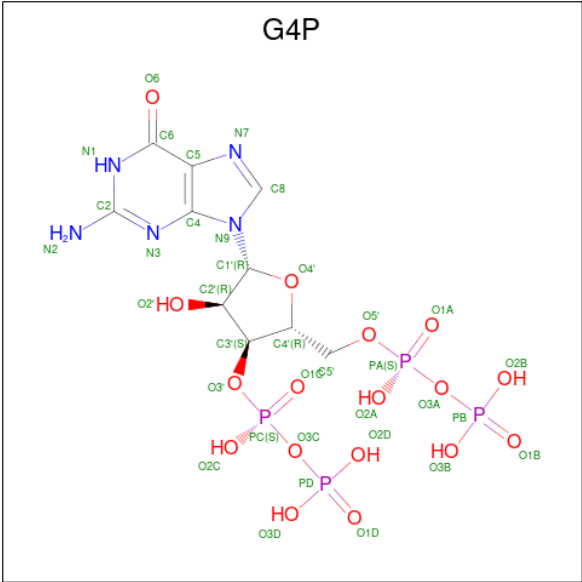
- Molecule 7 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	D	1	Total	Mg	0	0
			1	1		
7	J	1	Total	Mg	0	0
			1	1		

- Molecule 8 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	D	2	Total	Zn	0	0
			2	2		
8	J	2	Total	Zn	0	0
			2	2		
8	M	1	Total	Zn	0	0
			1	1		

- Molecule 9 is GUANOSINE-5',3'-TETRAPHOSPHATE (CCD ID: G4P) (formula: C₁₀H₁₇N₅O₁₇P₄).

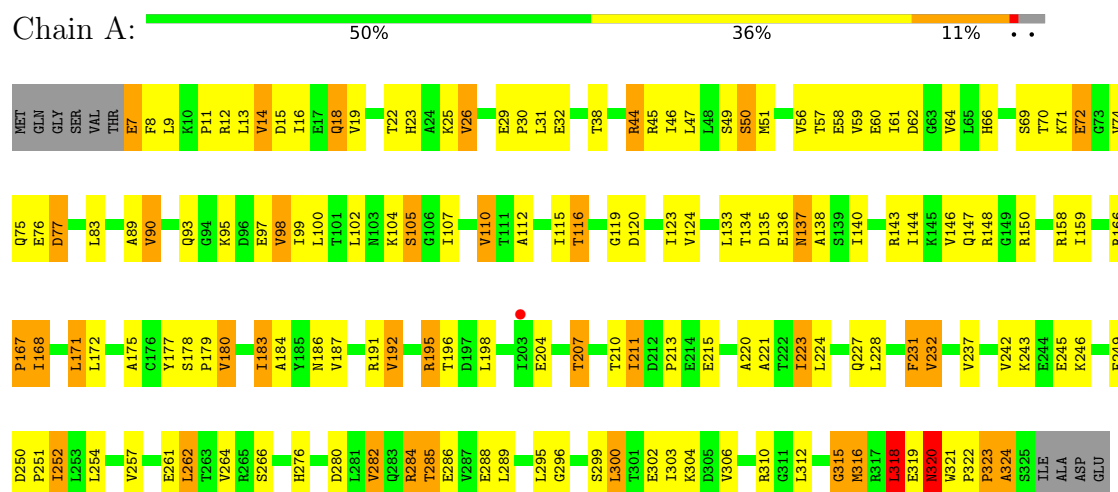


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
9	E	1	Total	C	N	O	P	0	0
			36	10	5	17	4		
9	J	1	Total	C	N	O	P	0	0
			36	10	5	17	4		
9	M	1	Total	C	N	O	P	0	0
			36	10	5	17	4		

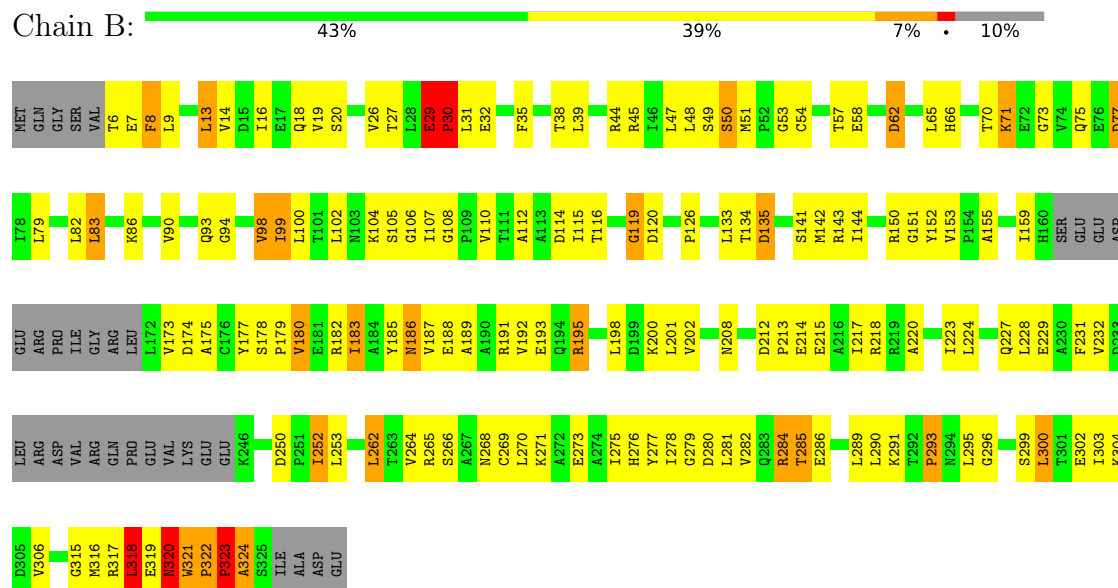
3 Residue-property plots

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.

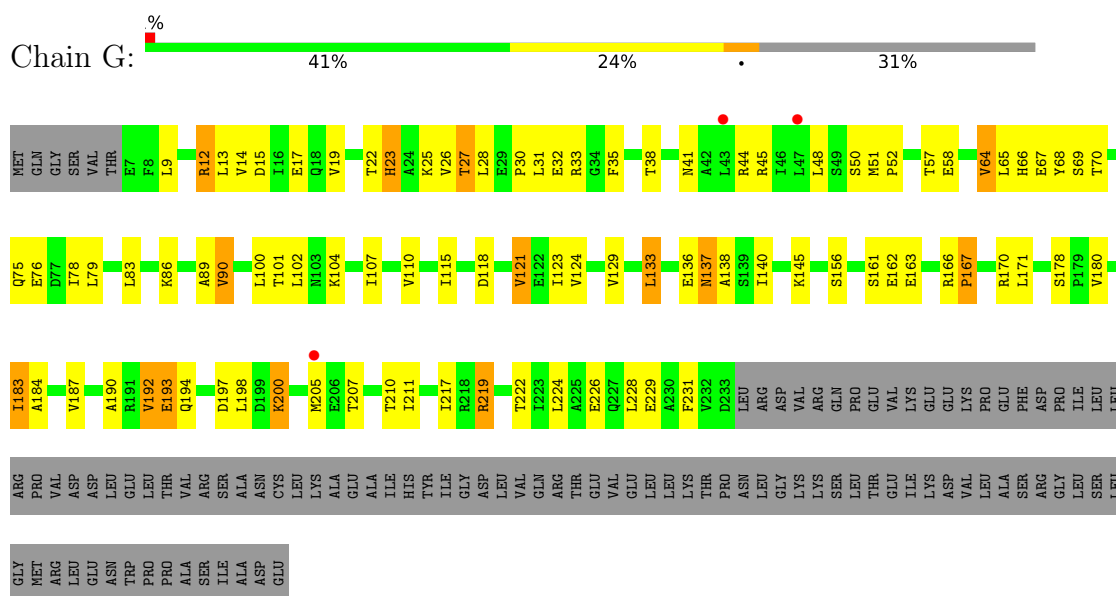
• Molecule 1: DNA-directed RNA polymerase subunit alpha



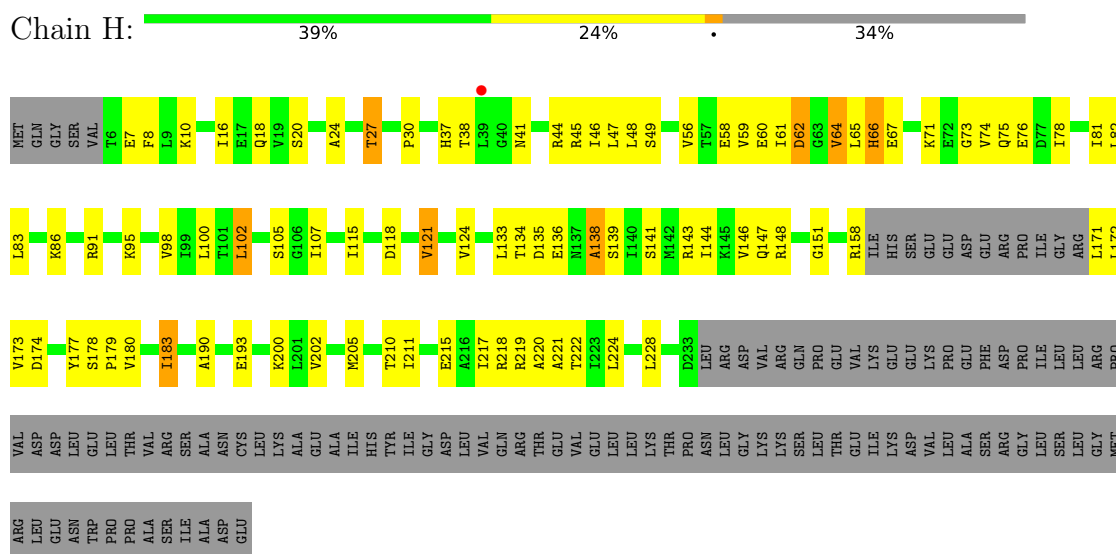
• Molecule 1: DNA-directed RNA polymerase subunit alpha



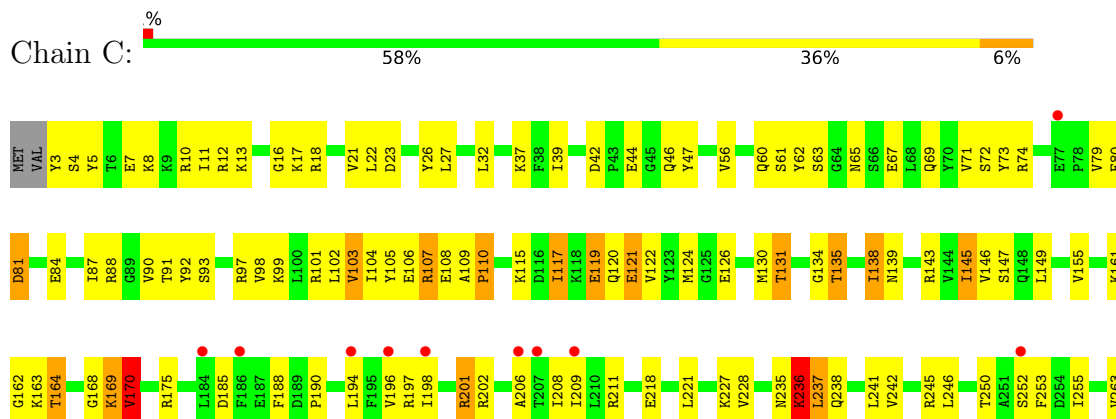
• Molecule 1: DNA-directed RNA polymerase subunit alpha

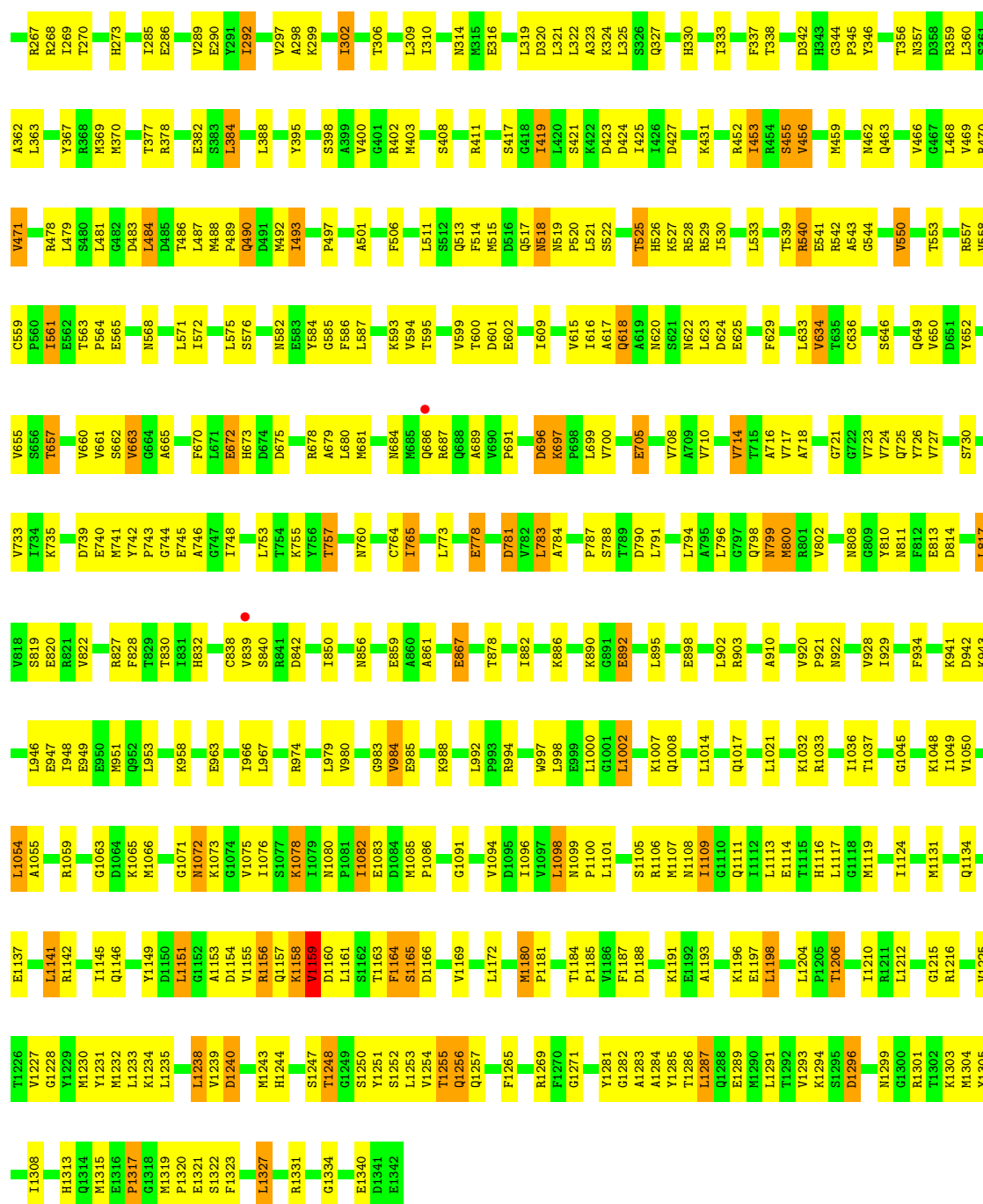


• Molecule 1: DNA-directed RNA polymerase subunit alpha

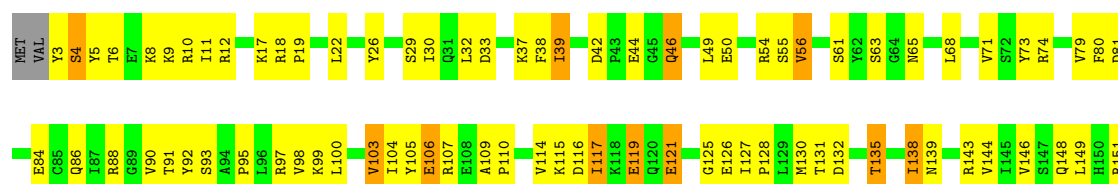


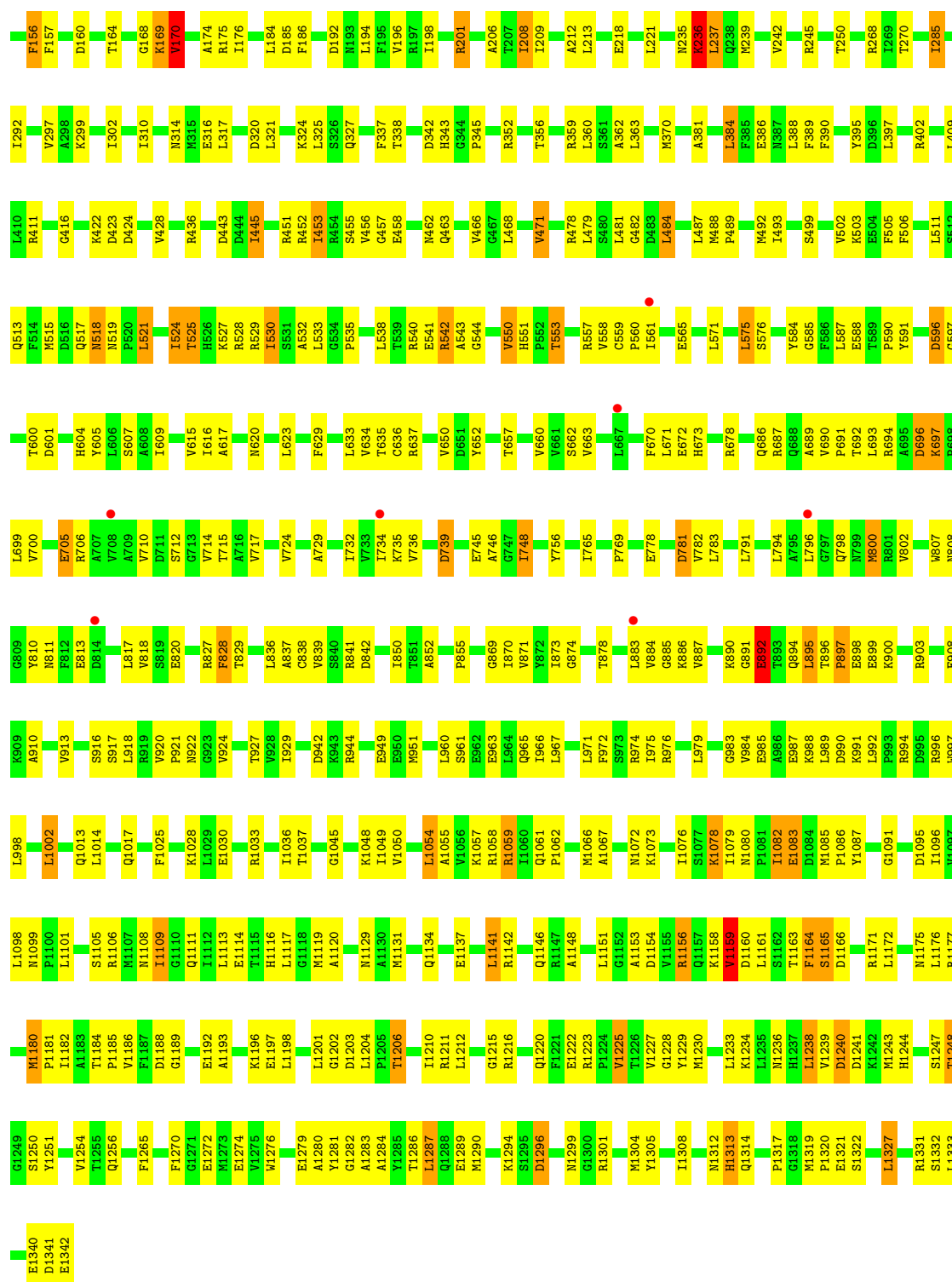
• Molecule 2: DNA-directed RNA polymerase subunit beta





• Molecule 2: DNA-directed RNA polymerase subunit beta

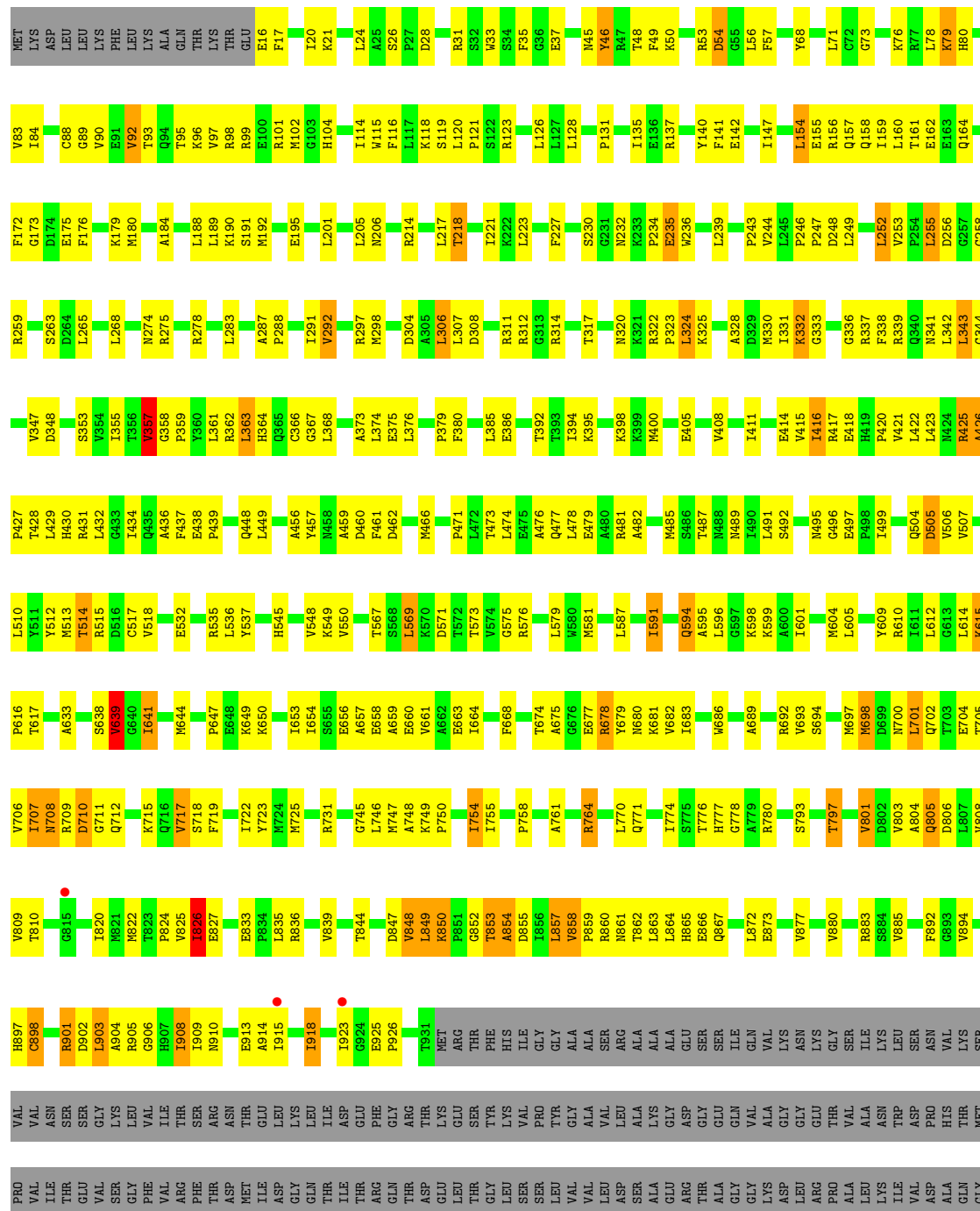


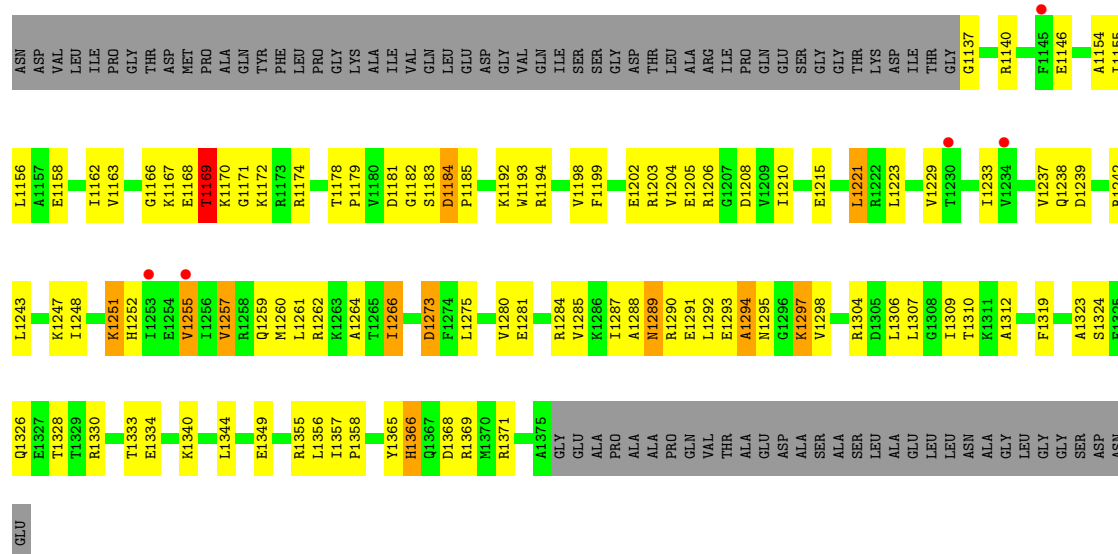


• Molecule 3: DNA-directed RNA polymerase subunit beta'



L1221	E146	ALA	HIS	V894	E811	V717	I624	R535	P451	K370	R275	L188	E91
R1222	A1147	GLN	THR	Y899	G815	S718	I624	L536	L452	L374	L279	L189	V92
H1227	R1148	GLY	MET	G900	R719	F719	F629	Y537	A459	E375		K190	T93
A1228	R1149	ASN	PRO	R901	H817	S721	F629	R538	D460	L376	L285	K191	Q94
V1229	P1153	VAL	ILE	D902	I820	I722	A633	L541	F461	F377	V292	M192	T95
T1230	A1154	LEU	THR	A904	R821	Y723	S638	A542	D462	K378		D193	K96
I1233	L1155	PRO	GLJ	R905	R822	M724	V639	H545	G463	P379	E295	L194	V97
V1234	L1156	GLY	VAL	G906	T823	M725	V639	H545	G463			L194	R98
I1237	I1159	THR	GLY	R907	R824	R731	I641	V548	Q382	Y382		C198	R99
Q1238	S1160	PHE	THR	I908	R825	G732	P647	K549	G383	P304		L201	E100
D1239	G1161	VAL	ILE	N910	E827	S733	P647	V550	L385	A305		R202	R101
Y1240	V1162	ARG	THR	E913	G828	A734	E658	R551	E386	L306		P110	M102
S1241	F1164	THR	PHE	L930	G829	Q736	A659	E556	T392	G310		S210	H113
L1243	F1165	TYR	ASP	A914	D830	I737	A659	E556	T392	R312		K213	I114
Q1244	G1166	PHE	NET	I915	R831	R738	V661	A559	I394	G313		K214	K118
K1247	K1167	LEU	ILE	I918	K832	G742	I664	A559	Q477	R314		K215	S119
I1248	E1168	PRO	ASP	S922	E833	M743	I664	E562	L478	M400		R220	L120
K1251	T1169	GLY	THR	S922	P834	R744	Q667	L563	A482	E405		R220	P121
H1252	G1171	LYS	GLN	E925	L835	G745	F668	V564	L483	A406		E225	R123
K1255	K1172	ILE	ILE	P926	V839	L746	L746	L569	M484	V407		E226	L126
I1256	R1173	VAL	THR	Q929	R842	A748	V673	K570	S486	V408		F227	L127
V1257	R1174	GLN	GLY	L930	T844	P750	T674	D571	L487	L324		S230	L128
L1261	T1177	ASP	THR	T931	T844	P750	T674	T572	M488	K321		E235	D129
I1266	D1181	LEU	GLY	R932	D847	I754	E677	T573	M489	V415		E236	M130
G1270	S1183	THR	GLY	R933	R849	P758	E677	R576	I490	R417		E237	R133
S1272	P1185	VAL	LEU	T934	L849	P758	E677	R576	I491	E418		I238	
D1273	E1187	ILE	THR	F935	K860	P758	E677	R576	I491	E418		E136	E136
F1274	E1188	GLY	GLY	R936	P851	M762	E677	R576	I491	E418		E137	R137
L1275	E1189	LEU	LEU	T937	G852	F763	E677	R576	I491	E418		E137	R137
Q1279	K1192	ASP	VAL	GLY	T853	R764	E677	R576	I491	E418		E137	R137
V1280	V1193	THR	VAL	ALA	L857	L770	E677	R576	I491	E418		E137	R137
E1281	R1194	LEU	VAL	ALA	V888	L770	E677	R576	I491	E418		E137	R137
Y1282	V1198	ARG	VAL	ALA	P859	I774	E677	R576	I491	E418		E137	R137
S1283	F1199	ILE	VAL	ALA	R860	S775	E677	R576	I491	E418		E137	R137
R1284	E1200	ILE	ASP	ALA	R861	T776	E677	R576	I491	E418		E137	R137
K1285	G1201	ALA	ASP	ALA	T862	H777	E677	R576	I491	E418		E137	R137
I1287	E1202	GLY	ALA	ALA	L863	L783	E677	R576	I491	E418		E137	R137
A1286	R1203	ASP	GLY	GLY	L864	L783	E677	R576	I491	E418		E137	R137
R1289	R1206	THR	THR	GLY	L872	T786	E677	R576	I491	E418		E137	R137
E1293	I1210	GLY	GLY	GLY	N875	K789	E677	R576	I491	E418		E137	R137
A1294	E1215	GLY	GLY	GLY	S876	T790	E677	R576	I491	E418		E137	R137
K1297	D1219	ASP	ASP	ASP	V877	S793	E677	R576	I491	E418		E137	R137
V1298	I1220	ASP	ASP	ASP	V877	S793	E677	R576	I491	E418		E137	R137
G1299		ASP	ASP	ASP	V877	S793	E677	R576	I491	E418		E137	R137
A1300		ASP	ASP	ASP	V877	S793	E677	R576	I491	E418		E137	R137





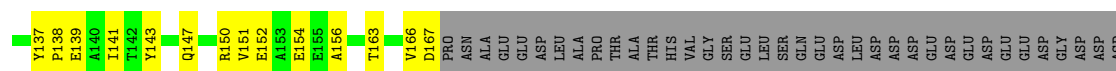
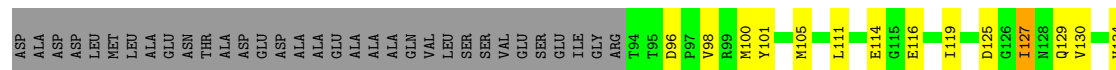
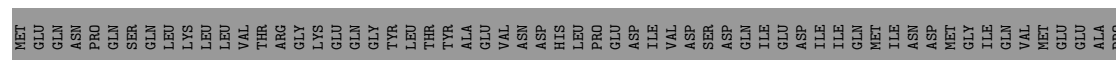
- Molecule 4: DNA-directed RNA polymerase subunit omega



- Molecule 4: DNA-directed RNA polymerase subunit omega



- Molecule 5: RNA polymerase sigma factor RpoD





D71	P72	V73	D74	R75	A76	A77	Q78	E79		F82	S83	L84	E85	L86	R87	N88	R89	D90	R91	E92	R93	K94		K97	K98	I99		K105	V106	E107	D108		F111	G112		I121	G122	I123	R124	R125	L126	E127	A128	R129		D133		I136		K139	T140	L141		I144	R145	E146	K147		G151
-----	-----	-----	-----	-----	-----	-----	-----	-----	--	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	--	-----	-----	-----	--	------	------	------	------	--	------	------	--	------	------	------	------	------	------	------	------	------	--	------	--	------	--	------	------	------	--	------	------	------	------	--	------

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	187.27Å 205.26Å 311.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.82 – 4.29 46.82 – 4.29	Depositor EDS
% Data completeness (in resolution range)	93.7 (46.82-4.29) 81.3 (46.82-4.29)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.85 (at 4.29Å)	Xtriage
Refinement program	PHENIX (1.11.1 _2575: ???)	Depositor
R, R_{free}	0.221 , 0.259 0.226 , 0.259	Depositor DCC
R_{free} test set	2000 reflections (2.43%)	wwPDB-VP
Wilson B-factor (Å ²)	187.9	Xtriage
Anisotropy	0.126	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 279.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	57643	wwPDB-VP
Average B, all atoms (Å ²)	279.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.91% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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, G4P, 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	A	0.24	0/2524	0.71	2/3421 (0.1%)
1	B	0.27	0/2326	0.98	6/3153 (0.2%)
1	G	0.19	0/1777	0.57	0/2408
1	H	0.22	0/1681	0.68	0/2278
2	C	0.21	0/10739	0.64	1/14489 (0.0%)
2	I	0.20	0/10735	0.61	1/14484 (0.0%)
3	D	0.22	0/9235	0.67	2/12472 (0.0%)
3	J	0.21	0/9140	0.67	2/12341 (0.0%)
4	E	0.20	0/693	0.56	0/935
4	K	0.22	0/629	0.73	1/847 (0.1%)
5	F	0.19	0/3864	0.61	2/5194 (0.0%)
5	L	0.19	0/3872	0.59	0/5205
6	M	0.32	0/1155	0.82	0/1549
All	All	0.22	0/58370	0.66	17/78776 (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
1	B	0	1

There are no bond length outliers.

The worst 5 of 17 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	29	GLU	CA-C-N	9.35	131.53	119.84
1	B	29	GLU	C-N-CA	9.35	131.53	119.84
4	K	3	ARG	N-CA-C	-7.96	93.41	107.73

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	236	LYS	N-CA-C	7.52	126.82	110.80
1	B	30	PRO	CA-C-N	-7.12	112.94	122.99

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	323	PRO	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	A	2490	0	2542	126	0
1	B	2297	0	2350	147	1
1	G	1755	0	1773	63	0
1	H	1662	0	1687	58	0
2	C	10570	0	10582	401	0
2	I	10566	0	10576	373	0
3	D	9095	0	9222	393	0
3	J	9001	0	9167	379	1
4	E	691	0	695	26	0
4	K	627	0	634	28	0
5	F	3813	0	3880	142	0
5	L	3821	0	3884	125	0
6	M	1140	0	1119	64	0
7	D	1	0	0	0	0
7	J	1	0	0	0	0
8	D	2	0	0	0	0
8	J	2	0	0	0	0
8	M	1	0	0	0	0
9	E	36	0	11	3	0
9	J	36	0	11	4	0
9	M	36	0	11	7	0
All	All	57643	0	58144	2101	1

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 18.

The worst 5 of 2101 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:B:29:GLU:HB3	1:B:30:PRO:CD	1.26	1.46
1:B:29:GLU:CB	1:B:30:PRO:HD2	1.35	1.45
1:B:279:GLY:HA3	1:B:321:TRP:CZ2	1.55	1.41
5:F:600:HIS:O	5:F:604:SER:HB3	1.33	1.29
1:B:253:LEU:O	1:B:321:TRP:HZ2	0.98	1.26

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:250:ASP:OD2	3:J:675:ALA:N[2_555]	2.03	0.17

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	317/329 (96%)	249 (78%)	53 (17%)	15 (5%)	2	16
1	B	291/329 (88%)	225 (77%)	45 (16%)	21 (7%)	1	11
1	G	225/329 (68%)	200 (89%)	21 (9%)	4 (2%)	6	33
1	H	212/329 (64%)	197 (93%)	11 (5%)	4 (2%)	6	32
2	C	1338/1342 (100%)	1201 (90%)	114 (8%)	23 (2%)	7	35
2	I	1338/1342 (100%)	1196 (89%)	119 (9%)	23 (2%)	7	35
3	D	1169/1407 (83%)	1042 (89%)	99 (8%)	28 (2%)	4	27
3	J	1151/1407 (82%)	1024 (89%)	100 (9%)	27 (2%)	5	28
4	E	87/91 (96%)	81 (93%)	4 (5%)	2 (2%)	5	28

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	K	77/91 (85%)	68 (88%)	5 (6%)	4 (5%)	1	15
5	F	462/613 (75%)	425 (92%)	28 (6%)	9 (2%)	6	32
5	L	463/613 (76%)	424 (92%)	32 (7%)	7 (2%)	8	38
6	M	138/151 (91%)	128 (93%)	7 (5%)	3 (2%)	5	29
All	All	7268/8373 (87%)	6460 (89%)	638 (9%)	170 (2%)	5	28

5 of 170 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	14	VAL
1	A	107	ILE
1	A	136	GLU
1	A	167	PRO
1	A	242	VAL

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	278/286 (97%)	224 (81%)	54 (19%)	1	8
1	B	256/286 (90%)	220 (86%)	36 (14%)	3	15
1	G	193/286 (68%)	162 (84%)	31 (16%)	2	13
1	H	183/286 (64%)	168 (92%)	15 (8%)	10	31
2	C	1155/1157 (100%)	1022 (88%)	133 (12%)	5	20
2	I	1154/1157 (100%)	1032 (89%)	122 (11%)	6	22
3	D	962/1168 (82%)	859 (89%)	103 (11%)	6	22
3	J	960/1168 (82%)	857 (89%)	103 (11%)	6	22
4	E	72/75 (96%)	64 (89%)	8 (11%)	6	21
4	K	67/75 (89%)	59 (88%)	8 (12%)	5	19
5	F	417/540 (77%)	375 (90%)	42 (10%)	7	24
5	L	418/540 (77%)	374 (90%)	44 (10%)	6	23

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
6	M	121/131 (92%)	110 (91%)	11 (9%)	9 28
All	All	6236/7155 (87%)	5526 (89%)	710 (11%)	5 20

5 of 710 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	I	484	LEU
3	J	376	LEU
2	I	561	ILE
2	I	481	LEU
2	I	1156	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 96 such sidechains are listed below:

Mol	Chain	Res	Type
2	I	133	ASN
2	I	1314	GLN
2	I	330	HIS
2	I	894	GLN
3	J	419	HIS

5.3.3 RNA [i](#)

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 7 are monoatomic - leaving 3 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	G4P	M	202	-	36,38,38	1.97	12 (33%)	56,61,61	1.41	8 (14%)
9	G4P	J	2004	-	36,38,38	1.94	13 (36%)	56,61,61	1.30	5 (8%)
9	G4P	E	101	-	36,38,38	1.85	12 (33%)	56,61,61	1.30	6 (10%)

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	G4P	M	202	-	-	8/27/43/43	0/3/3/3
9	G4P	J	2004	-	-	7/27/43/43	0/3/3/3
9	G4P	E	101	-	-	1/27/43/43	0/3/3/3

The worst 5 of 37 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	M	202	G4P	C5-N7	-5.12	1.28	1.39
9	J	2004	G4P	C5-N7	-4.93	1.29	1.39
9	E	101	G4P	C5-N7	-4.68	1.29	1.39
9	M	202	G4P	C6-N1	-4.05	1.31	1.38
9	J	2004	G4P	C6-N1	-3.89	1.31	1.38

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
9	M	202	G4P	C5-C4-N3	-4.92	120.56	128.39
9	E	101	G4P	C5-C4-N3	-4.20	121.71	128.39
9	J	2004	G4P	C5-C4-N3	-4.04	121.97	128.39
9	J	2004	G4P	N9-C8-N7	-3.85	106.25	113.40
9	M	202	G4P	N9-C8-N7	-3.80	106.35	113.40

There are no chirality outliers.

5 of 16 torsion outliers are listed below:

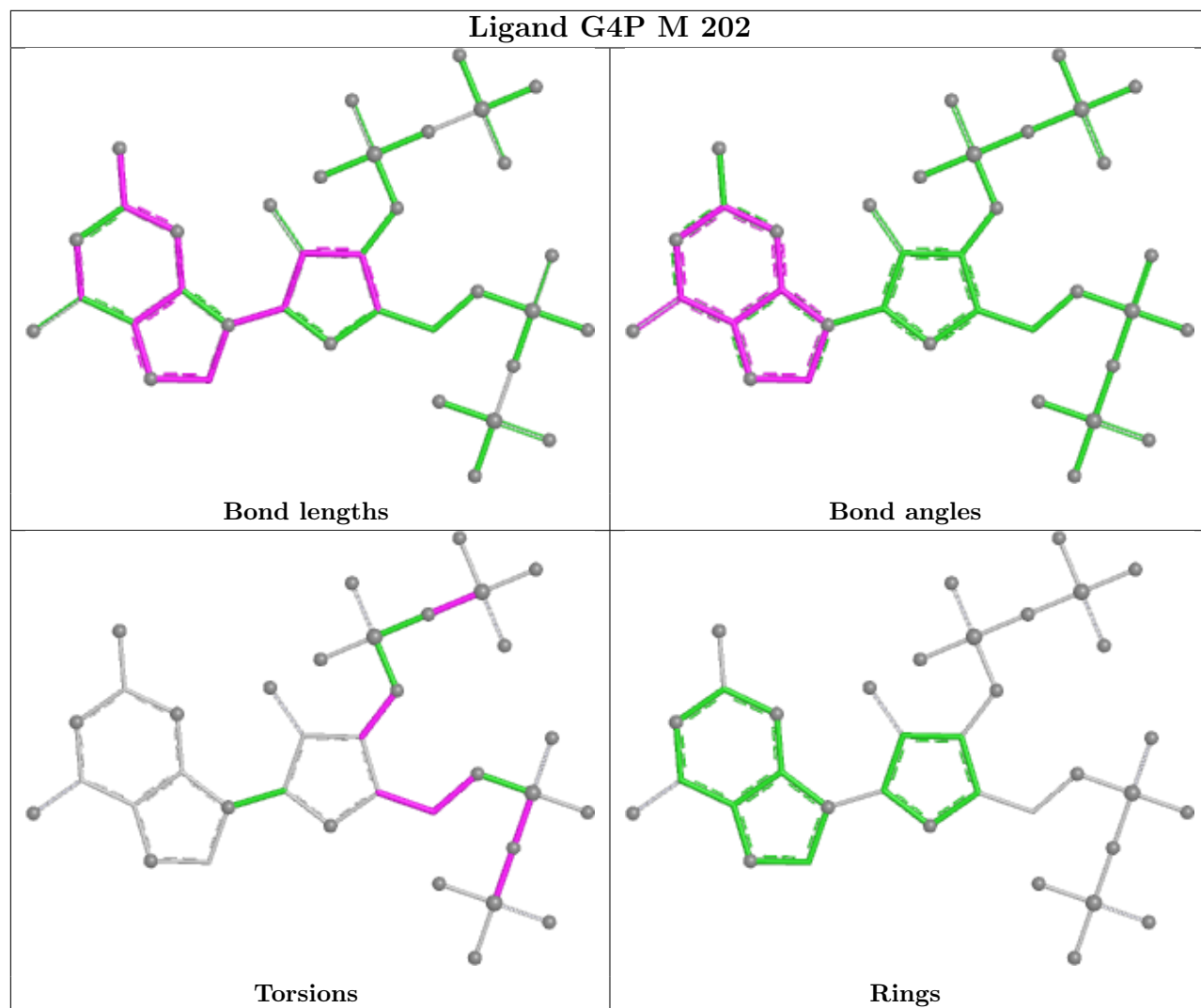
Mol	Chain	Res	Type	Atoms
9	J	2004	G4P	C3'-O3'-PC-O3C
9	M	202	G4P	PA-O3A-PB-O2B
9	M	202	G4P	C4'-C5'-O5'-PA
9	M	202	G4P	PB-O3A-PA-O5'
9	J	2004	G4P	C3'-O3'-PC-O2C

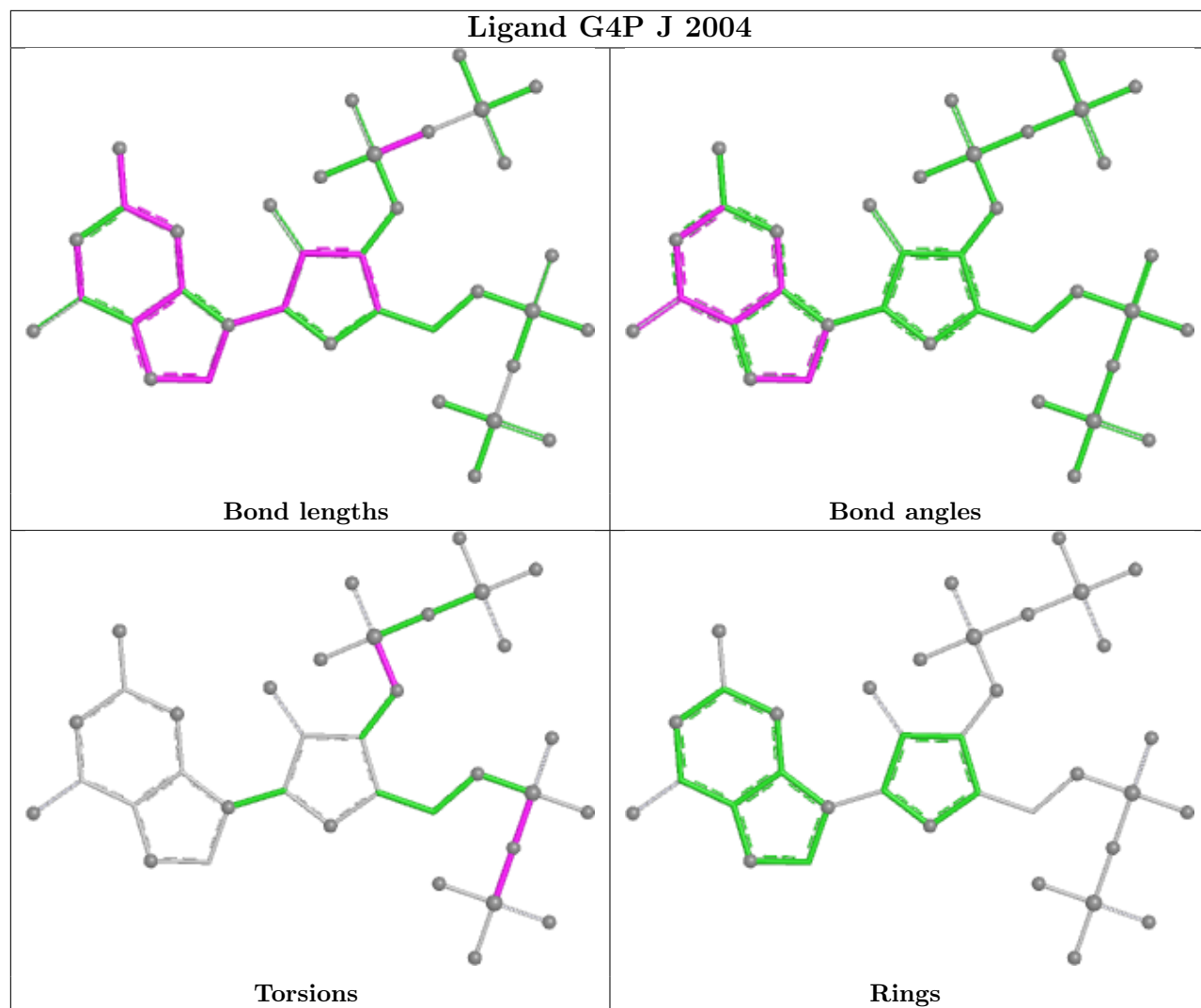
There are no ring outliers.

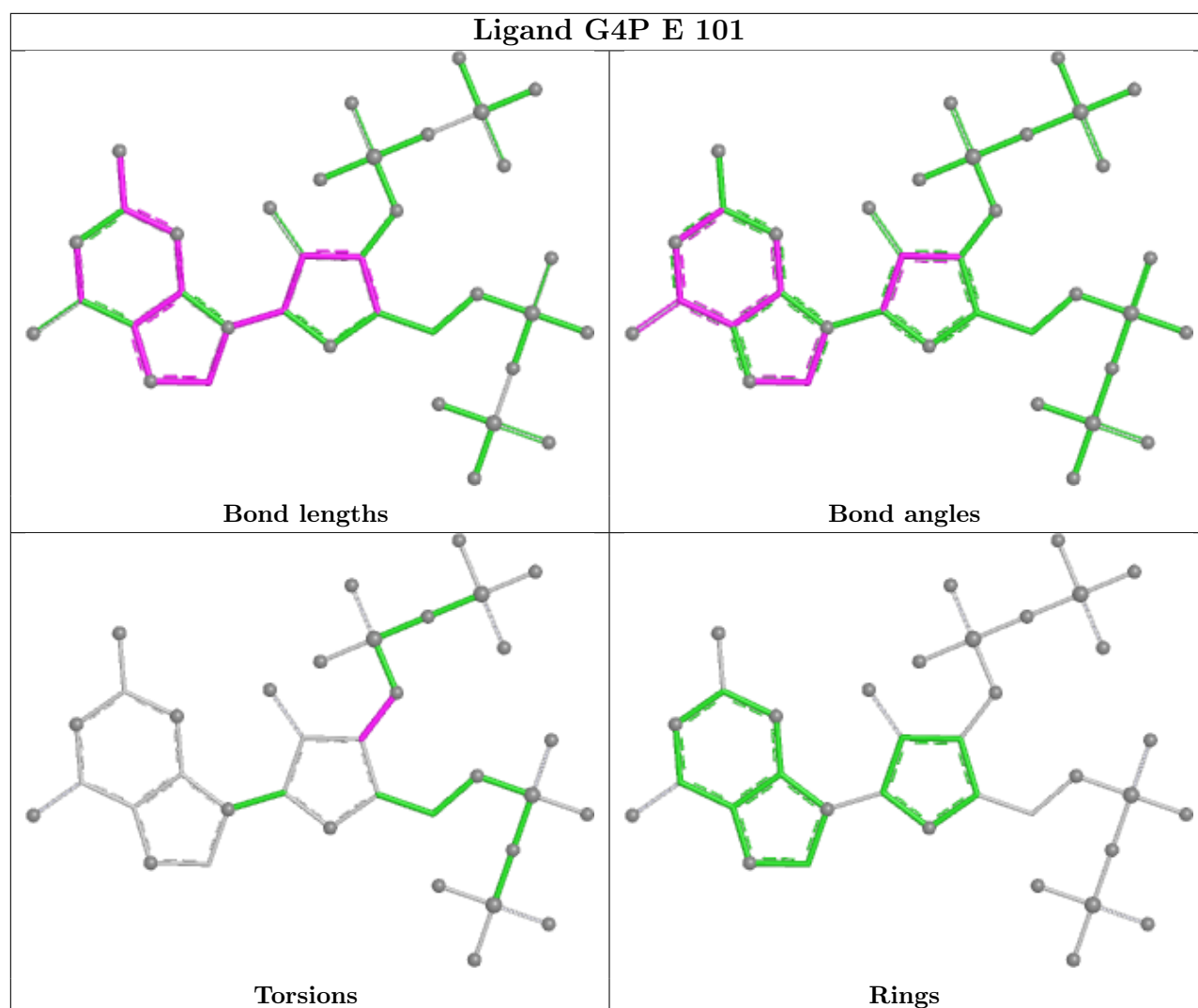
3 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	M	202	G4P	7	0
9	J	2004	G4P	4	0
9	E	101	G4P	3	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.

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(Å ²)	Q<0.9
1	A	319/329 (96%)	-0.43	1 (0%) 90 80	211, 260, 453, 474	0
1	B	297/329 (90%)	-0.42	0 100 100	204, 303, 444, 497	0
1	G	227/329 (68%)	-0.40	3 (1%) 75 59	257, 320, 369, 399	0
1	H	216/329 (65%)	-0.35	1 (0%) 87 75	288, 336, 378, 393	0
2	C	1340/1342 (99%)	-0.48	12 (0%) 81 66	154, 240, 335, 399	0
2	I	1340/1342 (99%)	-0.50	7 (0%) 87 75	200, 281, 382, 521	0
3	D	1173/1407 (83%)	-0.37	9 (0%) 82 68	152, 223, 328, 395	0
3	J	1155/1407 (82%)	-0.45	8 (0%) 84 71	197, 262, 344, 388	0
4	E	89/91 (97%)	-0.56	1 (1%) 78 62	202, 266, 330, 365	0
4	K	79/91 (86%)	-0.60	0 100 100	336, 411, 518, 552	0
5	F	468/613 (76%)	-0.52	0 100 100	188, 293, 445, 515	0
5	L	469/613 (76%)	-0.62	1 (0%) 91 83	223, 313, 427, 446	0
6	M	140/151 (92%)	-0.28	0 100 100	317, 361, 378, 386	0
All	All	7312/8373 (87%)	-0.46	43 (0%) 85 73	152, 271, 395, 552	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	482	ALA	4.2
1	G	205	MET	4.0
3	J	1230	THR	4.0
3	J	1255	VAL	3.4
1	G	43	LEU	3.3

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

6.4 Ligands [i](#)

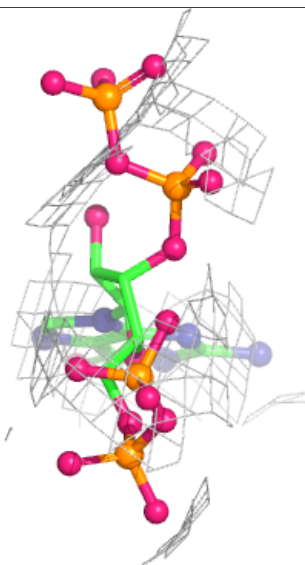
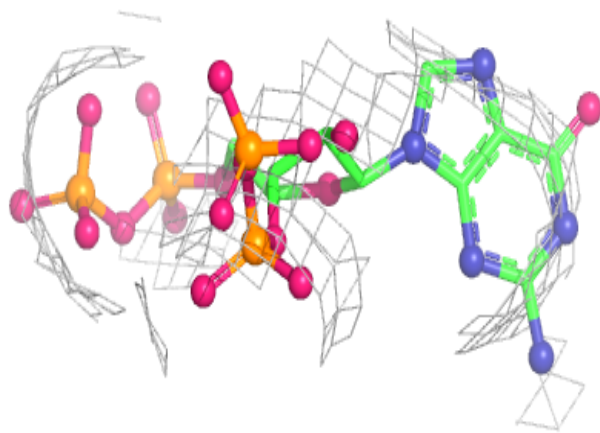
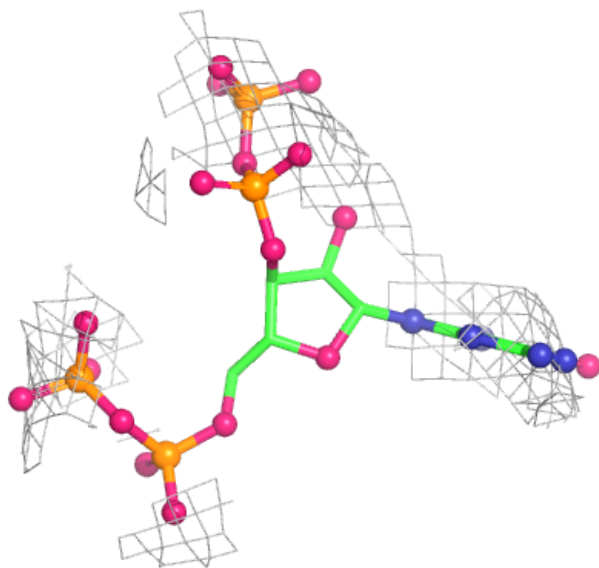
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors($\text{\AA}^2$)	Q<0.9
9	G4P	J	2004	36/36	0.71	0.07	306,370,455,463	0
9	G4P	E	101	36/36	0.91	0.06	199,253,286,301	0
9	G4P	M	202	36/36	0.91	0.05	267,326,403,408	0
8	ZN	M	201	1/1	0.93	0.05	662,662,662,662	0
7	MG	J	2001	1/1	0.94	0.06	184,184,184,184	0
7	MG	D	2001	1/1	0.96	0.04	162,162,162,162	0
8	ZN	J	2002	1/1	0.98	0.03	264,264,264,264	0
8	ZN	D	2002	1/1	1.00	0.04	273,273,273,273	0
8	ZN	J	2003	1/1	1.00	0.07	176,176,176,176	0
8	ZN	D	2003	1/1	1.00	0.12	295,295,295,295	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.

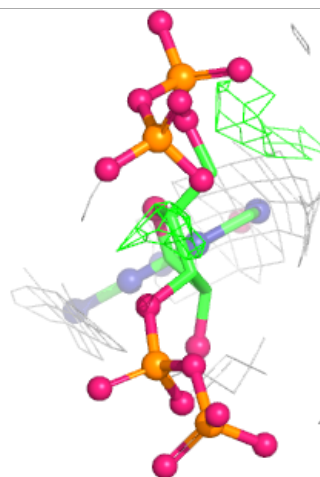
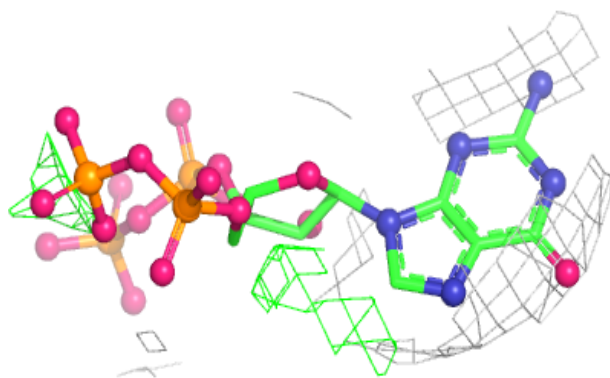
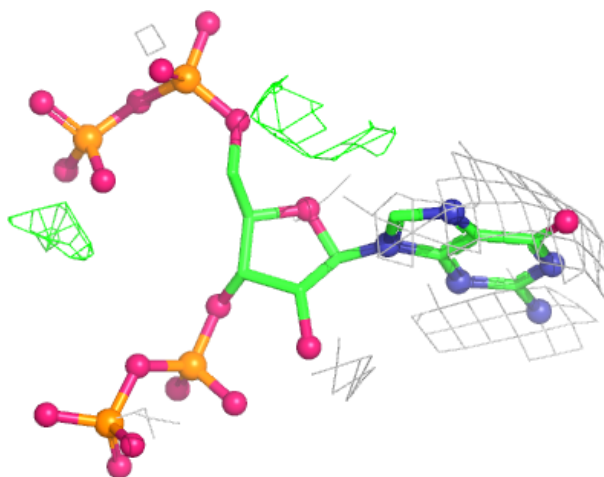
Electron density around G4P J 2004:

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



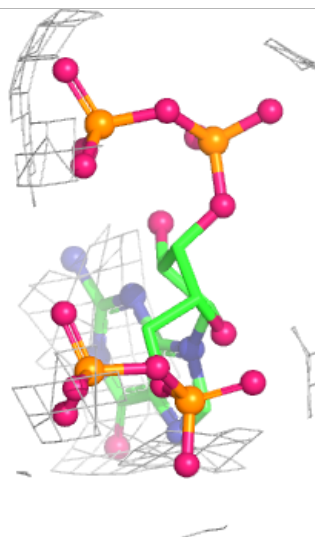
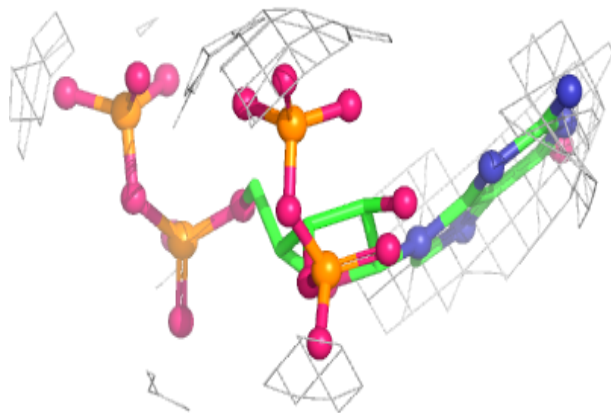
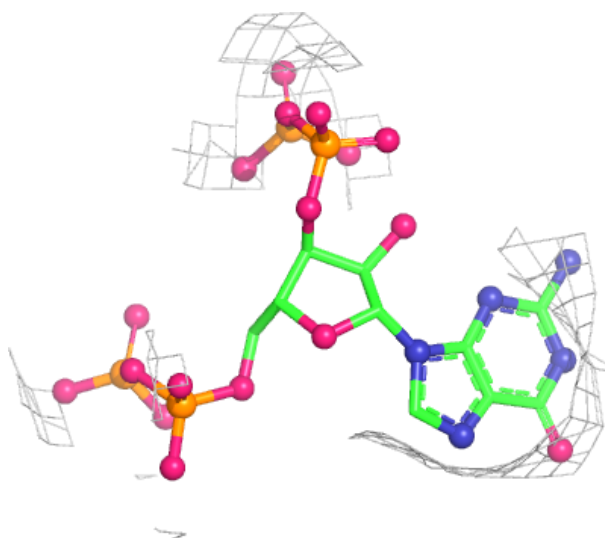
Electron density around G4P E 101:

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



Electron density around G4P M 202:

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



6.5 Other polymers [i](#)

There are no such residues in this entry.