



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 14, 2026 – 10:12 AM UTC

PDB ID : 2E2I / pdb\_00002e2i  
Title : RNA polymerase II elongation complex in 5 mM Mg+2 with 2'-dGTP  
Authors : Wang, D.; Bushnell, D.A.; Westover, K.D.; Kaplan, C.D.; Kornberg, R.D.  
Deposited on : 2006-11-14  
Resolution : 3.41 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<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)  
Xtrriage (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

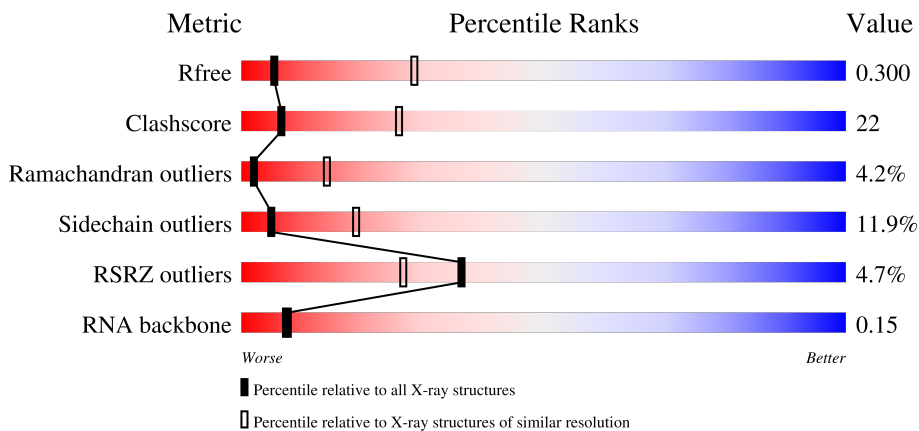
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


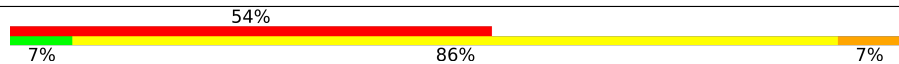
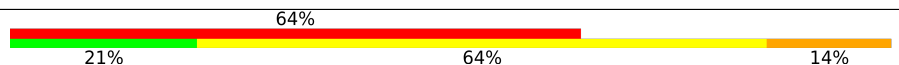

The reported resolution of this entry is 3.41 Å.

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	1210 (3.48-3.36)
Clashscore	190562	1234 (3.48-3.36)
Ramachandran outliers	187476	1222 (3.48-3.36)
Sidechain outliers	187428	1222 (3.48-3.36)
RSRZ outliers	180081	1210 (3.48-3.36)
RNA backbone	3983	1162 (3.84-3.00)

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  $\geq 3$ , 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	R	10	
2	T	28	
3	N	14	
4	A	1733	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	B	1224	<p>3% 50% 33% 8% • 7%</p>
6	C	318	<p>51% 26% 7% 16%</p>
7	E	215	<p>6% 73% 22% • •</p>
8	F	155	<p>% 37% 17% •• 45%</p>
9	H	146	<p>10% 60% 24% 7% • 8%</p>
10	I	122	<p>2% 70% 25% ••</p>
11	J	70	<p>% 53% 27% 11% • 7%</p>
12	K	120	<p>% 56% 33% 6% 5%</p>
13	L	70	<p>9% 44% 14% 6% • 34%</p>

## 2 Entry composition [i](#)

There are 16 unique types of molecules in this entry. The entry contains 29722 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 5'-R(\*AP\*UP\*CP\*GP\*AP\*GP\*AP\*GP\*GP\*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	R	10	216	98	45	64	9	0	0	0

- Molecule 2 is a DNA chain called 28-MER DNA template strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	T	28	564	270	102	165	27	0	0	0

- Molecule 3 is a DNA chain called 5'-D(\*CP\*TP\*GP\*CP\*TP\*TP\*AP\*TP\*CP\*GP\*GP\*TP\*AP\*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	N	14	284	137	49	85	13	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	A	1411	11098	6989	1944	2104	61	0	0	0

- Molecule 5 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
5	B	1143	9092	5753	1595	1688	56	0	0	0

- Molecule 6 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

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

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
7	E	213	1744	1107	308	318	11	0	0	0

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	F	86	697	445	118	131	3	0	0	0

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
9	H	134	1076	678	181	212	5	0	0	0

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit 9.

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

- Molecule 11 is a protein called DNA-directed RNA polymerases I/II/III subunit 10.

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

- Molecule 12 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

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

- Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide.

tide.

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

- Molecule 14 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

- Molecule 15 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

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

- Molecule 16 is 2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE (CCD ID: DGT) (formula: C<sub>10</sub>H<sub>16</sub>N<sub>5</sub>O<sub>13</sub>P<sub>3</sub>).



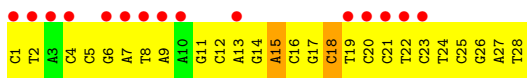
### 3 Residue-property plots [i](#)

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



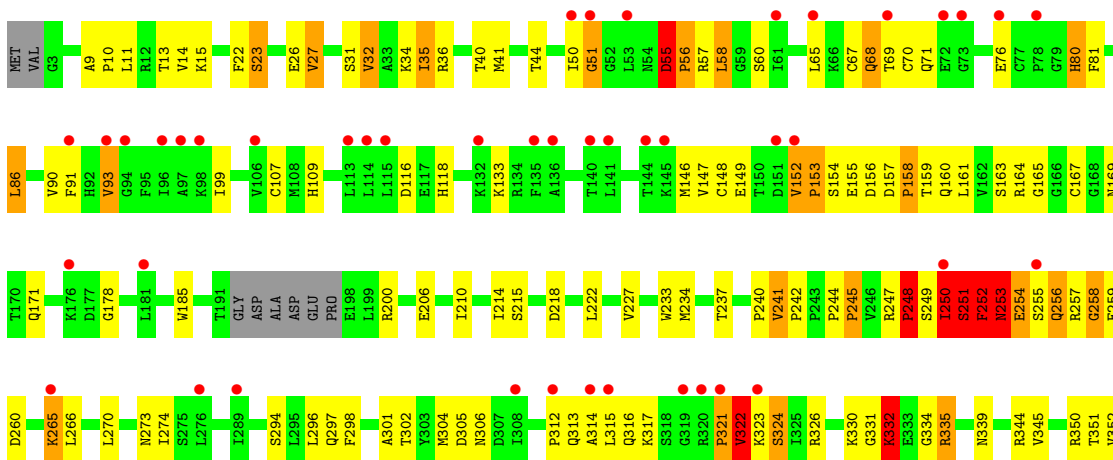
- Molecule 2: 28-MER DNA template strand



- Molecule 3: 5'-D(\*CP\*TP\*GP\*CP\*TP\*TP\*AP\*TP\*CP\*GP\*GP\*TP\*AP\*G)-3'

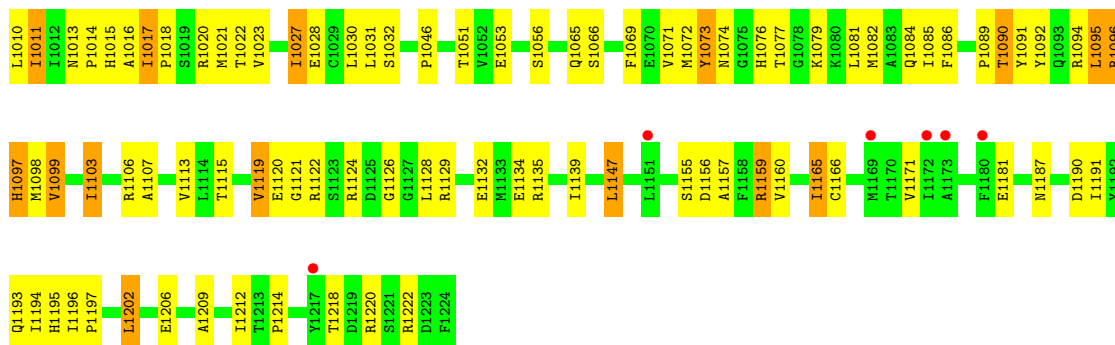


- Molecule 4: DNA-directed RNA polymerase II largest subunit



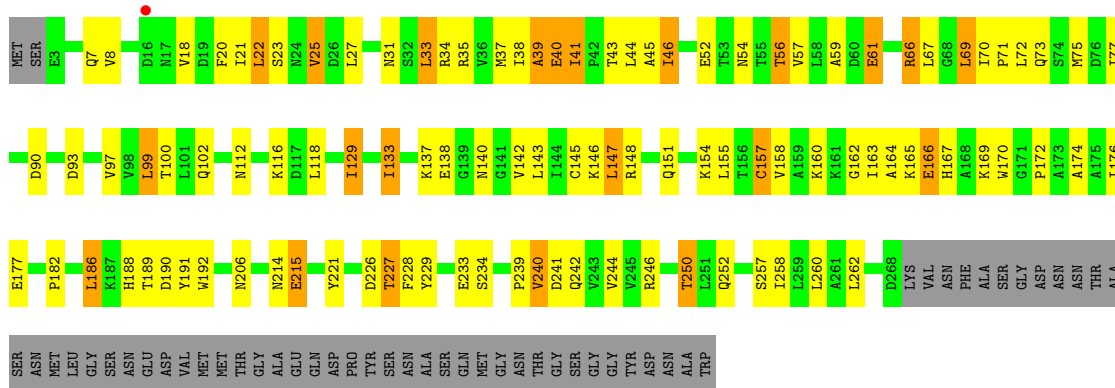
TYR	GLY	ASP	S1383	V1304	Y897	L808	T596	V612	D438	I353
SER	ALA	GLY	V1384	V1305	R898	T809	L897	SS13	H439	P357
PRO	TYR	PHE	T1385	L1306	W899	T899	L997	SS19	D440	N358
THR	GLY	ASP	E1307	E1307	D900	D900	L998	NS17	L443	L359
PRO	GLU	Q1187	T1308	T1308	L901	L901	R1001	NS18	F444	E360
SER	ALA	L1193	M1312	M1106	L902	L902	G1002	NS19	L361	D362
TYR	THR	R1194	L1313	M1111	T907	T907	R1001	NS20	D363	V364
PRO	THR	D1198	L1314	M1112	L908	L908	R1001	NS21	R446	G365
PRO	TYR	R1199	T1315	K1112	L909	L909	R1001	NS22	Q447	V366
PRO	GLY	R1199	E1315	T1113	D909	D909	R1001	NS23	R447	P367
THR	PRO	A1200	V1316	T1114	L913	L913	R1001	NS24	L450	
SER	PHE	A1201	F1316	S1115	L914	L914	R1001	NS25	S454	
PRO	GLY	M1202	D1323	L1116	F914	F914	R1001	NS26	M455	
SER	VAL	L1324	F1324	T1117	S915	S915	R1001	NS27	M456	
TYR	SER	K1205	R1325	V1118	T919	T919	R1001	NS28	A457	
SER	LEU	R1326	R1326	Y1119	W829	W829	R1001	NS29	H458	
PRO	VAL	I1327	I1327	Y1120	R830	R830	R1001	NS30	R459	
PRO	GLY	Y1328	Y1328	L1120	T831	T831	R1001	NS31	V460	
THR	GLY	F1220	T1329	H1124	T834	T834	R1001	NS32	T461	
SER	PHE	L1224	T1329	H1124	G835	G835	R1001	NS33	L463	
PRO	ASP	F1225	I1333	A1125	Y836	Y836	R1001	NS34	P464	
SER	PRO	I1225	D1334	A1126	R840	R840	R1001	NS35	Y465	
TYR	THR	D1334	D1334	D1127	R840	R840	R1001	NS36	S466	
SER	VAL	I1408	I1335	Q1128	M849	M849	R1001	NS37	L467	
PRO	VAL	L1409	M1336	E1129	R857	R857	R1001	NS38	L391	
PRO	THR	F1410	E1337	Q1130	M858	M858	R1001	NS39	V392	
SER	THR		V1338	Q1130	T867	T867	R1001	NS40	L393	
PRO	TYR		L1339	Q1130	Y868	Y868	R1001	NS41	L394	
PRO	LEU		G1340	Q1130	D871	D871	R1001	NS42	L472	
PRO	ALA		I1341	Q1130	M873	M873	R1001	NS43	S473	
PRO	ALA		I1341	Q1130	D874	D874	R1001	NS44	V474	
PRO	ALA		R1345	Q1130	A875	A875	R1001	NS45	T475	
PRO	TYR		A1347	Q1130	K880	K880	R1001	NS46	Y476	
PRO	ASP		L1348	Q1130	D884	D884	R1001	NS47	S477	
PRO	SER		Y1349	Q1130	Y885	Y885	R1001	NS48	F478	
PRO	GLY		K1350	Q1130	W954	W954	R1001	NS49	F482	
PRO	SER		E1351	Q1130	P955	P955	R1001	NS50	D483	
PRO	ASN		V1352	Q1130	L956	L956	R1001	NS51	G484	
PRO	ASP		I1353	Q1130	P957	P957	R1001	NS52	D485	
PRO	ALA		M1354	Q1130	Y958	Y958	R1001	NS53	E486	
PRO	ALA		N1354	Q1130	W959	W959	R1001	NS54	M487	
PRO	MET		V1355	Q1130	L960	L960	R1001	NS55	V491	
PRO	ALA		I1356	Q1130	T971	T971	R1001	NS56	E496	
PRO	ALA		I1356	Q1130	E771	E771	R1001	NS57	E499	
PRO	GLY		D1359	Q1130	K772	K772	R1001	NS58	E500	
PRO	THR		V1363	Q1130	R774	R774	R1001	NS59	L501	
PRO	TYR		M1364	Q1130	I775	I775	R1001	NS60	C505	
PRO	TYR		Y1365	Q1130	G778	G778	R1001	NS61	A506	
PRO	TYR		R1366	Q1130	T783	T783	R1001	NS62	V507	
PRO	TYR		H1367	Q1130	S793	S793	R1001	NS63	P508	
PRO	TYR		M1368	Q1130	P794	P794	R1001	NS64	L509	
PRO	PRO		A1369	Q1130	Y799	Y799	R1001	NS65	S510	
PRO	THR		L1370	Q1130	G798	G798	R1001	NS66	L511	
PRO	GLY		L1371	Q1130	F799	F799	R1001	NS67		
PRO	GLY		V1372	Q1130	V800	V800	R1001	NS68		
PRO	ALA		D1373	Q1130	S803	S803	R1001	NS69		
PRO	THR		T1376	Q1130	Y804	Y804	R1001	NS70		
PRO	PRO		T1382	Q1130	L805	L805	R1001	NS71		
PRO	SER			Q1130			R1001	NS72		
PRO	SER			Q1130			R1001	NS73		
PRO	SER			Q1130			R1001	NS74		
PRO	SER			Q1130			R1001	NS75		
PRO	SER			Q1130			R1001	NS76		
PRO	SER			Q1130			R1001	NS77		
PRO	SER			Q1130			R1001	NS78		
PRO	SER			Q1130			R1001	NS79		
PRO	SER			Q1130			R1001	NS80		
PRO	SER			Q1130			R1001	NS81		
PRO	SER			Q1130			R1001	NS82		
PRO	SER			Q1130			R1001	NS83		
PRO	SER			Q1130			R1001	NS84		
PRO	SER			Q1130			R1001	NS85		
PRO	SER			Q1130			R1001	NS86		
PRO	SER			Q1130			R1001	NS87		
PRO	SER			Q1130			R1001	NS88		
PRO	SER			Q1130			R1001	NS89		
PRO	SER			Q1130			R1001	NS90		
PRO	SER			Q1130			R1001	NS91		
PRO	SER			Q1130			R1001	NS92		
PRO	SER			Q1130			R1001	NS93		
PRO	SER			Q1130			R1001	NS94		
PRO	SER			Q1130			R1001	NS95		
PRO	SER			Q1130			R1001	NS96		
PRO	SER			Q1130			R1001	NS97		
PRO	SER			Q1130			R1001	NS98		
PRO	SER			Q1130			R1001	NS99		
PRO	SER			Q1130			R1001	NS100		





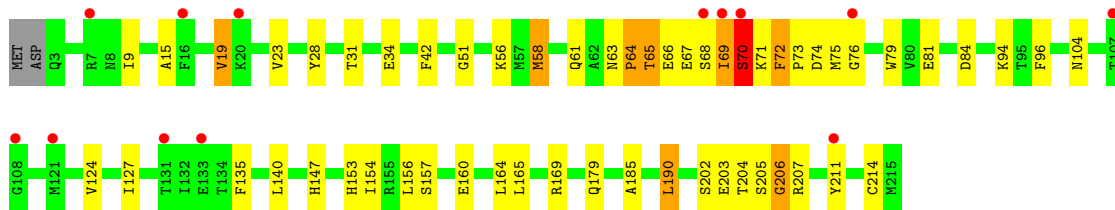
- Molecule 6: DNA-directed RNA polymerase II 45 kDa polypeptide

Chain C: 51% 26% 7% 16%



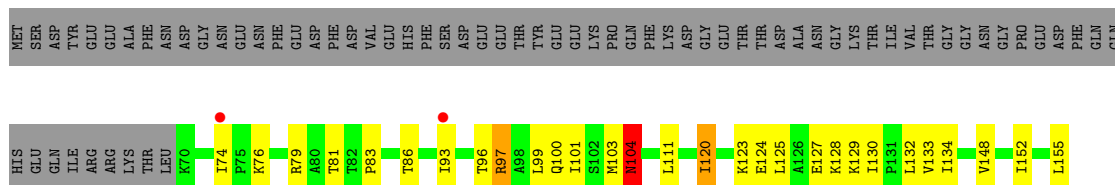
- Molecule 7: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide

Chain E: 6% 73% 22% ..



- Molecule 8: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

Chain F: % 37% 17% .. 45%



- Molecule 9: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	168.68Å 223.52Å 193.94Å 90.00° 100.54° 90.00°	Depositor
Resolution (Å)	50.00 – 3.41 50.00 – 3.41	Depositor EDS
% Data completeness (in resolution range)	98.0 (50.00-3.41) 98.0 (50.00-3.41)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.14	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.01 (at 3.40Å)	Xtrriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.266 , 0.316 0.257 , 0.300	Depositor DCC
$R_{free}$ test set	4732 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	92.7	Xtrriage
Anisotropy	0.155	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 81.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	29722	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	116.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.72% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality i

### 5.1 Standard geometry i

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

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	0.74	0/243	1.42	6/378 (1.6%)
2	T	0.37	0/631	1.07	9/970 (0.9%)
3	N	0.31	0/317	0.86	2/488 (0.4%)
4	A	0.82	4/11294 (0.0%)	1.07	32/15270 (0.2%)
5	B	0.99	9/9268 (0.1%)	1.19	47/12496 (0.4%)
6	C	0.87	1/2133 (0.0%)	1.09	5/2891 (0.2%)
7	E	0.64	0/1780	0.92	2/2395 (0.1%)
8	F	0.70	0/709	0.92	0/956
9	H	0.67	1/1094 (0.1%)	0.89	3/1480 (0.2%)
10	I	0.75	1/989 (0.1%)	0.90	2/1331 (0.2%)
11	J	1.00	1/541 (0.2%)	1.21	1/727 (0.1%)
12	K	0.82	1/937 (0.1%)	1.03	0/1265
13	L	1.04	0/365	1.17	0/485
All	All	0.86	18/30301 (0.1%)	1.09	109/41132 (0.3%)

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	7
5	B	0	11
7	E	0	1
11	J	0	2
All	All	0	21

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	507	VAL	CA-CB	8.35	1.59	1.53

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	846	ILE	CA-CB	-6.34	1.46	1.54
11	J	10	CYS	CA-CB	6.21	1.62	1.53
5	B	1027	ILE	CA-CB	-6.12	1.46	1.54
5	B	882	THR	CA-CB	6.09	1.62	1.53

The worst 5 of 109 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	B	635	ARG	CA-C-N	11.71	134.48	119.84
5	B	635	ARG	C-N-CA	11.71	134.48	119.84
1	R	9	G	C5'-C4'-C3'	-10.64	100.04	116.00
2	T	18	DC	C4'-C3'-O3'	9.94	124.90	110.00
11	J	9	SER	N-CA-C	8.83	120.91	111.28

There are no chirality outliers.

5 of 21 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	248	PRO	Peptide
4	A	250	ILE	Peptide
4	A	251	SER	Peptide
4	A	252	PHE	Peptide
4	A	254	GLU	Peptide

## 5.2 Too-close contacts [i](#)

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	216	0	109	13	0
2	T	564	0	316	88	0
3	N	284	0	161	19	0
4	A	11098	0	11174	450	0
5	B	9092	0	9135	520	0
6	C	2095	0	2051	96	0
7	E	1744	0	1772	47	0
8	F	697	0	720	20	0
9	H	1076	0	1052	35	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	I	971	0	928	17	0
11	J	532	0	542	51	0
12	K	919	0	929	40	0
13	L	363	0	388	6	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	62	0	24	2	0
All	All	29722	0	29301	1276	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 22.

The worst 5 of 1276 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:69:LEU:HG	5:B:70:ILE:CG2	1.32	1.55
2:T:19:DT:H2'	2:T:20:DC:C6	1.41	1.51
5:B:142:VAL:CG1	5:B:144:GLY:HA3	1.43	1.45
5:B:142:VAL:HG13	5:B:144:GLY:CA	1.57	1.34
5:B:69:LEU:CA	5:B:70:ILE:HB	1.58	1.28

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	
4	A	1401/1733 (81%)	1186 (85%)	160 (11%)	55 (4%)	2	15
5	B	1129/1224 (92%)	938 (83%)	133 (12%)	58 (5%)	1	11
6	C	264/318 (83%)	226 (86%)	33 (12%)	5 (2%)	6	26
7	E	211/215 (98%)	176 (83%)	29 (14%)	6 (3%)	4	19
8	F	84/155 (54%)	74 (88%)	7 (8%)	3 (4%)	2	16
9	H	130/146 (89%)	104 (80%)	21 (16%)	5 (4%)	2	15
10	I	117/122 (96%)	95 (81%)	19 (16%)	3 (3%)	4	21
11	J	63/70 (90%)	55 (87%)	4 (6%)	4 (6%)	1	7
12	K	112/120 (93%)	102 (91%)	8 (7%)	2 (2%)	6	26
13	L	44/70 (63%)	30 (68%)	7 (16%)	7 (16%)	0	0
All	All	3555/4173 (85%)	2986 (84%)	421 (12%)	148 (4%)	2	14

5 of 148 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	50	ILE
4	A	55	ASP
4	A	56	PRO
4	A	250	ILE
4	A	312	PRO

### 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	
4	A	1233/1520 (81%)	1098 (89%)	135 (11%)	6	22
5	B	991/1061 (93%)	855 (86%)	136 (14%)	3	14
6	C	234/274 (85%)	203 (87%)	31 (13%)	4	15
7	E	195/197 (99%)	183 (94%)	12 (6%)	16	42
8	F	76/137 (56%)	69 (91%)	7 (9%)	8	29
9	H	118/128 (92%)	104 (88%)	14 (12%)	5	19

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
10	I	113/116 (97%)	104 (92%)	9 (8%)	11	35
11	J	60/65 (92%)	52 (87%)	8 (13%)	4	15
12	K	99/102 (97%)	82 (83%)	17 (17%)	2	9
13	L	40/57 (70%)	34 (85%)	6 (15%)	3	11
All	All	3159/3657 (86%)	2784 (88%)	375 (12%)	5	19

5 of 375 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	906	SER
6	C	138	GLU
5	B	964	VAL
5	B	1124	ARG
6	C	250	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 86 such sidechains are listed below:

Mol	Chain	Res	Type
5	B	842	ASN
6	C	140	ASN
5	B	975	GLN
5	B	1097	HIS
7	E	147	HIS

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	9/10 (90%)	5 (55%)	1 (11%)

All (5) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	2	U
1	R	3	C
1	R	7	A
1	R	8	G
1	R	9	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	R	1	A

## 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, 9 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
16	DGT	B	1308[A]	-	32,33,33	1.71	7 (21%)	48,52,52	1.69	9 (18%)
16	DGT	B	1308[B]	-	32,33,33	1.49	5 (15%)	48,52,52	1.66	6 (12%)

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
16	DGT	B	1308[A]	-	-	6/22/34/34	0/3/3/3
16	DGT	B	1308[B]	-	-	6/22/34/34	0/3/3/3

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	B	1308[A]	DGT	PA-O3A	4.68	1.64	1.59
16	B	1308[A]	DGT	PB-O3B	4.16	1.64	1.59
16	B	1308[A]	DGT	PB-O3A	3.76	1.63	1.59
16	B	1308[B]	DGT	PB-O3B	3.70	1.63	1.59
16	B	1308[B]	DGT	PB-O3A	3.45	1.63	1.59

The worst 5 of 15 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
16	B	1308[B]	DGT	C5-C4-N3	-5.22	120.09	128.39
16	B	1308[A]	DGT	C5-C4-N3	-5.16	120.18	128.39
16	B	1308[B]	DGT	C2-N3-C4	4.61	120.24	112.30
16	B	1308[A]	DGT	C2-N3-C4	4.60	120.23	112.30
16	B	1308[A]	DGT	N9-C4-N3	4.14	134.23	125.95

There are no chirality outliers.

5 of 12 torsion outliers are listed below:

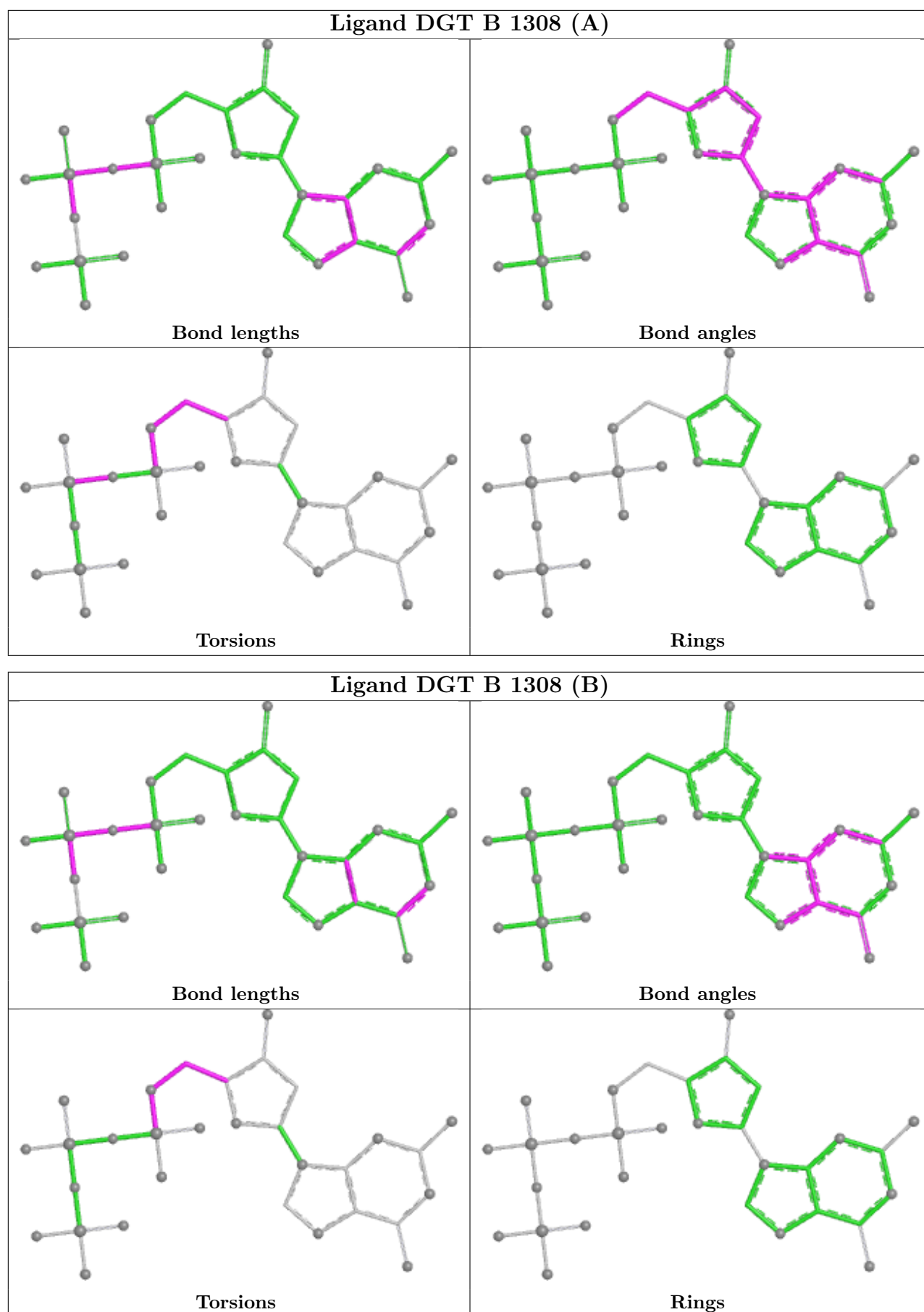
Mol	Chain	Res	Type	Atoms
16	B	1308[A]	DGT	C5'-O5'-PA-O3A
16	B	1308[A]	DGT	C5'-O5'-PA-O2A
16	B	1308[A]	DGT	C4'-C5'-O5'-PA
16	B	1308[B]	DGT	C5'-O5'-PA-O1A
16	B	1308[B]	DGT	C3'-C4'-C5'-O5'

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	B	1308[A]	DGT	2	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	R	10/10 (100%)	0.50	1 (10%) 12 12	72, 93, 202, 234	0
2	T	28/28 (100%)	1.97	15 (53%) 0 0	118, 195, 427, 447	0
3	N	14/14 (100%)	2.53	9 (64%) 0 0	270, 338, 424, 426	0
4	A	1411/1733 (81%)	0.45	74 (5%) 33 24	56, 105, 206, 247	0
5	B	1143/1224 (93%)	0.24	31 (2%) 56 41	27, 88, 156, 185	0
6	C	266/318 (83%)	0.06	1 (0%) 88 81	59, 91, 135, 151	0
7	E	213/215 (99%)	0.79	13 (6%) 27 21	96, 142, 228, 233	0
8	F	86/155 (55%)	0.42	2 (2%) 61 46	83, 117, 137, 141	0
9	H	134/146 (91%)	0.89	14 (10%) 11 12	35, 140, 171, 175	0
10	I	119/122 (97%)	0.52	3 (2%) 58 43	85, 121, 156, 165	0
11	J	65/70 (92%)	-0.04	1 (1%) 72 58	60, 82, 110, 117	0
12	K	114/120 (95%)	0.05	1 (0%) 81 69	77, 102, 119, 122	0
13	L	46/70 (65%)	1.08	6 (13%) 7 8	87, 165, 176, 179	0
All	All	3649/4225 (86%)	0.40	171 (4%) 36 27	27, 103, 198, 447	0

The worst 5 of 171 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
13	L	26	THR	5.1
4	A	72	GLU	4.7
5	B	89	GLU	4.3
4	A	320	ARG	4.2
9	H	1	MET	4.2

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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](#)

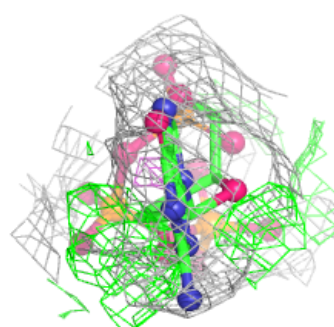
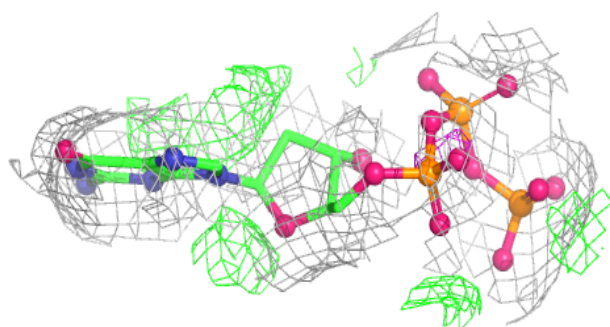
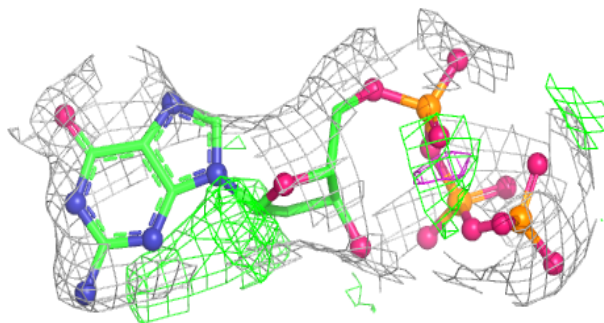
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
16	DGT	B	1308[A]	31/31	0.74	0.18	126,132,152,153	31
16	DGT	B	1308[B]	31/31	0.74	0.18	80,86,113,113	31
15	MG	A	2001	1/1	0.76	0.25	116,116,116,116	0
14	ZN	B	1307	1/1	0.96	0.06	143,143,143,143	0
14	ZN	A	1734	1/1	0.97	0.04	226,226,226,226	0
14	ZN	L	105	1/1	0.98	0.04	190,190,190,190	0
14	ZN	A	1735	1/1	0.98	0.04	139,139,139,139	0
14	ZN	I	204	1/1	0.99	0.02	106,106,106,106	0
14	ZN	J	101	1/1	0.99	0.04	105,105,105,105	0
14	ZN	I	203	1/1	0.99	0.07	131,131,131,131	0
14	ZN	C	319	1/1	1.00	0.01	95,95,95,95	0

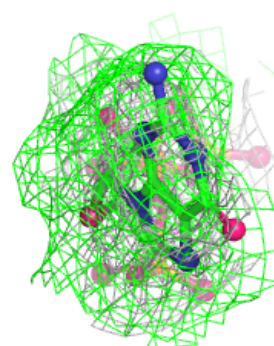
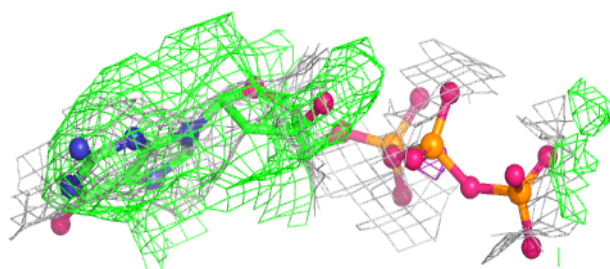
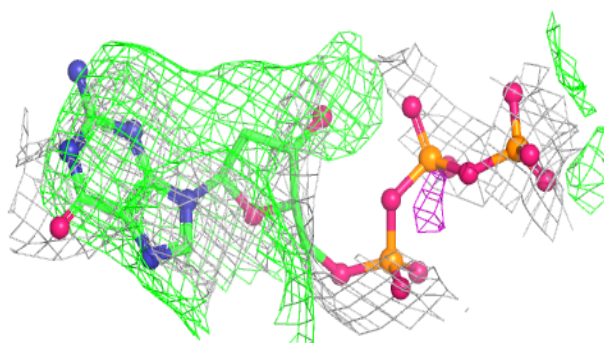
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

**Electron density around DGT B 1308 (A):**

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

**Electron density around DGT B 1308 (B):**

$2mF_o-DF_c$  (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.