



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 27, 2020 – 02:09 PM BST

PDB ID : 6UU6
Title : E. coli sigma-S transcription initiation complex with a 4-nt RNA and a UTP ("Old" crystal soaked with UTP, ddCTP, and dinucleotide ApG for 30 minutes)
Authors : Zuo, Y.; De, S.; Steitz, T.A.
Deposited on : 2019-10-30
Resolution : 4.20 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 following versions of software and data (see [references ⓘ](#)) 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 : 2.13
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) : 2.13

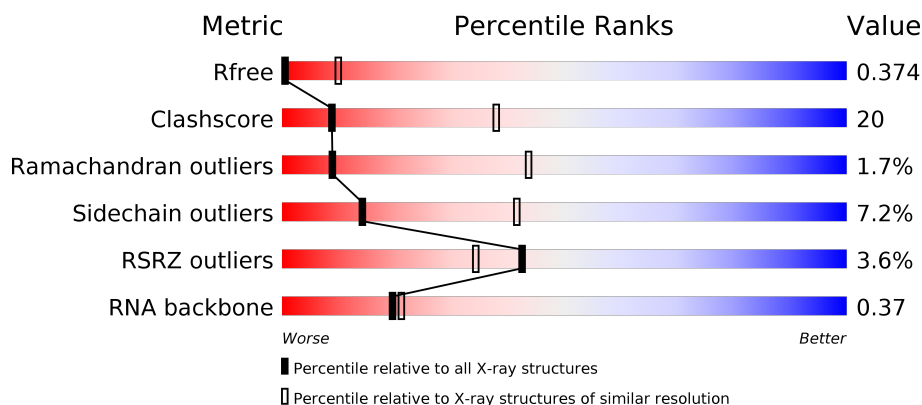
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.20 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}	130704	1005 (4.62-3.78)
Clashscore	141614	1044 (4.60-3.80)
Ramachandran outliers	138981	1000 (4.60-3.80)
Sidechain outliers	138945	1007 (4.62-3.78)
RSRZ outliers	127900	1063 (4.70-3.70)
RNA backbone	3102	1049 (5.04-3.00)

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 on 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	AAA	242	<div> <div>3%</div> <div>58% 33% 5%</div> </div>
1	BBB	242	<div> <div>5%</div> <div>59% 32% 6%</div> </div>
2	CCC	1342	<div> <div>3%</div> <div>63% 33%</div> </div>
3	DDD	1407	<div> <div>4%</div> <div>60% 34%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	EEE	90	
5	FFF	336	
6	111	50	
7	222	50	
8	333	4	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
10	ZN	DDD	1502	-	-	X	-
12	DOC	333	101	-	-	X	-
9	UTP	CCC	1401	-	-	X	-

2 Entry composition [i](#)

There are 12 unique types of molecules in this entry. The entry contains 29026 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	AAA	230	Total	C	N	O	S	0	0	0
			1787	1112	317	352	6			
1	BBB	228	Total	C	N	O	S	0	0	0
			1767	1100	312	349	6			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	-6	ALA	-	expression tag	UNP P0A7Z4
AAA	-5	HIS	-	expression tag	UNP P0A7Z4
AAA	-4	HIS	-	expression tag	UNP P0A7Z4
AAA	-3	HIS	-	expression tag	UNP P0A7Z4
AAA	-2	HIS	-	expression tag	UNP P0A7Z4
AAA	-1	HIS	-	expression tag	UNP P0A7Z4
AAA	0	HIS	-	expression tag	UNP P0A7Z4
BBB	-6	ALA	-	expression tag	UNP P0A7Z4
BBB	-5	HIS	-	expression tag	UNP P0A7Z4
BBB	-4	HIS	-	expression tag	UNP P0A7Z4
BBB	-3	HIS	-	expression tag	UNP P0A7Z4
BBB	-2	HIS	-	expression tag	UNP P0A7Z4
BBB	-1	HIS	-	expression tag	UNP P0A7Z4
BBB	0	HIS	-	expression tag	UNP P0A7Z4

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	CCC	1341	Total	C	N	O	S	0	0	0
			10577	6636	1842	2056	43			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	DDD	1362	Total	C	N	O	S	0	0	0
			10568	6633	1887	1998	50			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	EEE	79	Total	C	N	O	S	0	0	0
			627	382	118	126	1			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	FFF	277	Total	C	N	O	S	0	0	0
			2253	1411	415	423	4			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
FFF	2	GLY	SER	conflict	UNP P13445
FFF	33	GLU	GLN	conflict	UNP P13445
FFF	329	LEU	-	expression tag	UNP P13445
FFF	330	GLU	-	expression tag	UNP P13445
FFF	331	HIS	-	expression tag	UNP P13445
FFF	332	HIS	-	expression tag	UNP P13445
FFF	333	HIS	-	expression tag	UNP P13445
FFF	334	HIS	-	expression tag	UNP P13445
FFF	335	HIS	-	expression tag	UNP P13445
FFF	336	HIS	-	expression tag	UNP P13445

- Molecule 6 is a DNA chain called Synthetic DNA 50-mer (promoter non-template strand).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	111	30	Total	C	N	O	P	0	0	0
			618	294	111	183	30			

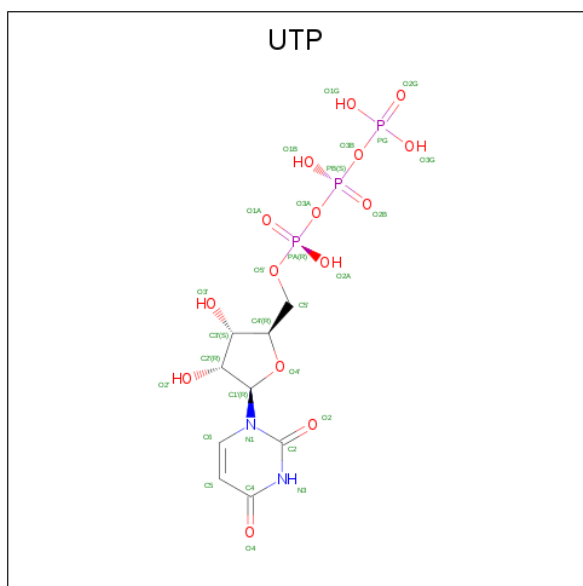
- Molecule 7 is a DNA chain called Synthetic DNA 50-mer (promoter template strand).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	222	35	Total	C	N	O	P	0	0	0
			716	342	132	208	34			

- Molecule 8 is a RNA chain called RNA 4-mer (dinucleotide ApG primed synthesis).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	333	3	Total	C	N	O	P	0	0	0
			62	29	12	19	2			

- Molecule 9 is URIDINE 5'-TRIPHOSPHATE (three-letter code: UTP) (formula: $C_9H_{15}N_2O_{15}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
9	CCC	1	Total	C	N	O	P	0	0
			29	9	2	15	3		

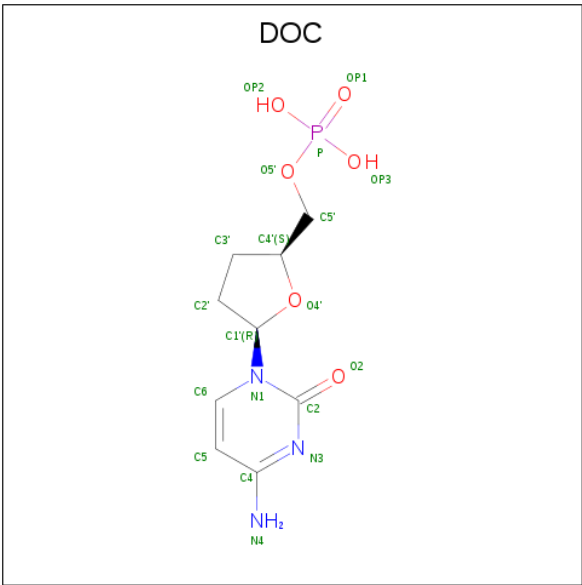
- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	DDD	2	Total	Zn	0	0
			2	2		

- Molecule 11 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	DDD	2	Total	Mg	0	0
			2	2		

- Molecule 12 is 2',3'-DIDEOXYCYTIDINE-5'-MONOPHOSPHATE (three-letter code: DOC) (formula: $C_9H_{14}N_3O_6P$).

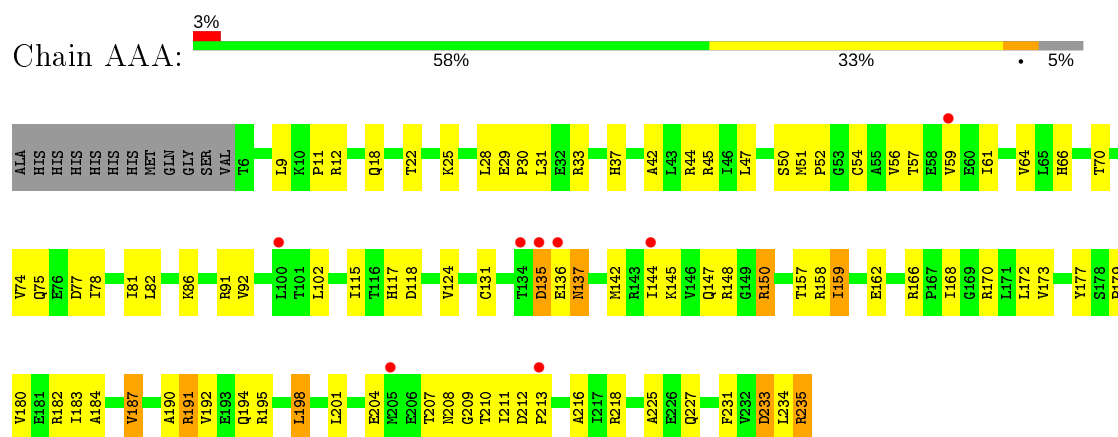


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
12	333	1	18	9	3	5	1	0	0

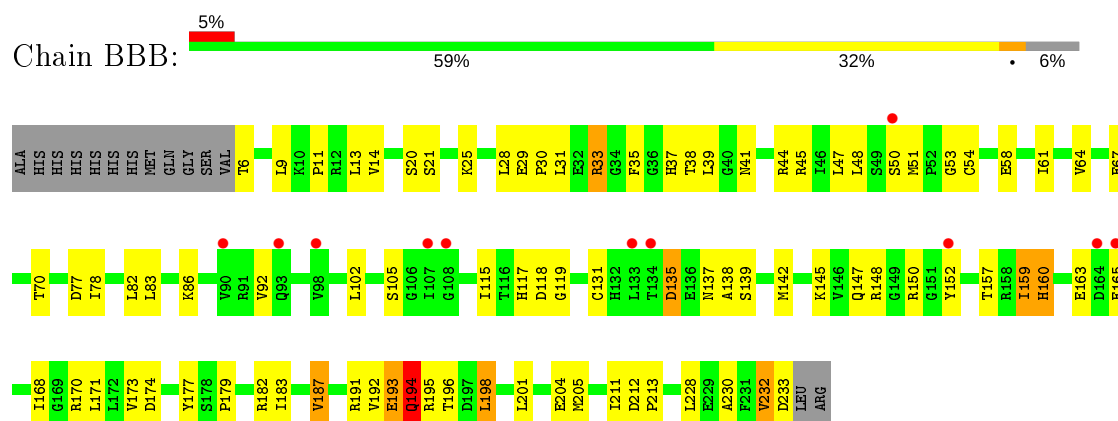
3 Residue-property plots [i](#)

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

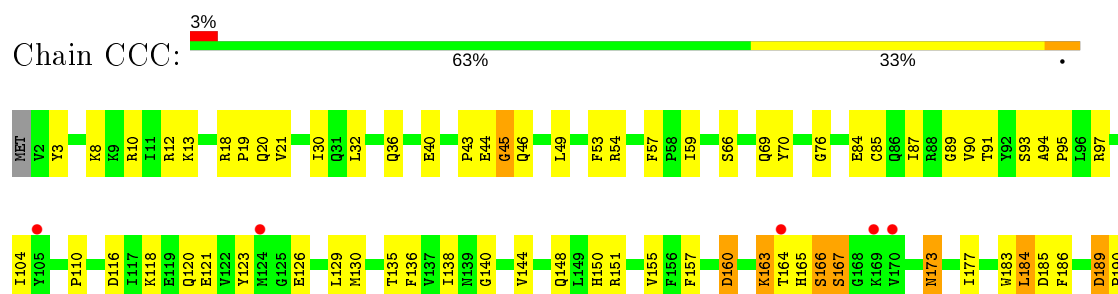
- Molecule 1: DNA-directed RNA polymerase subunit alpha

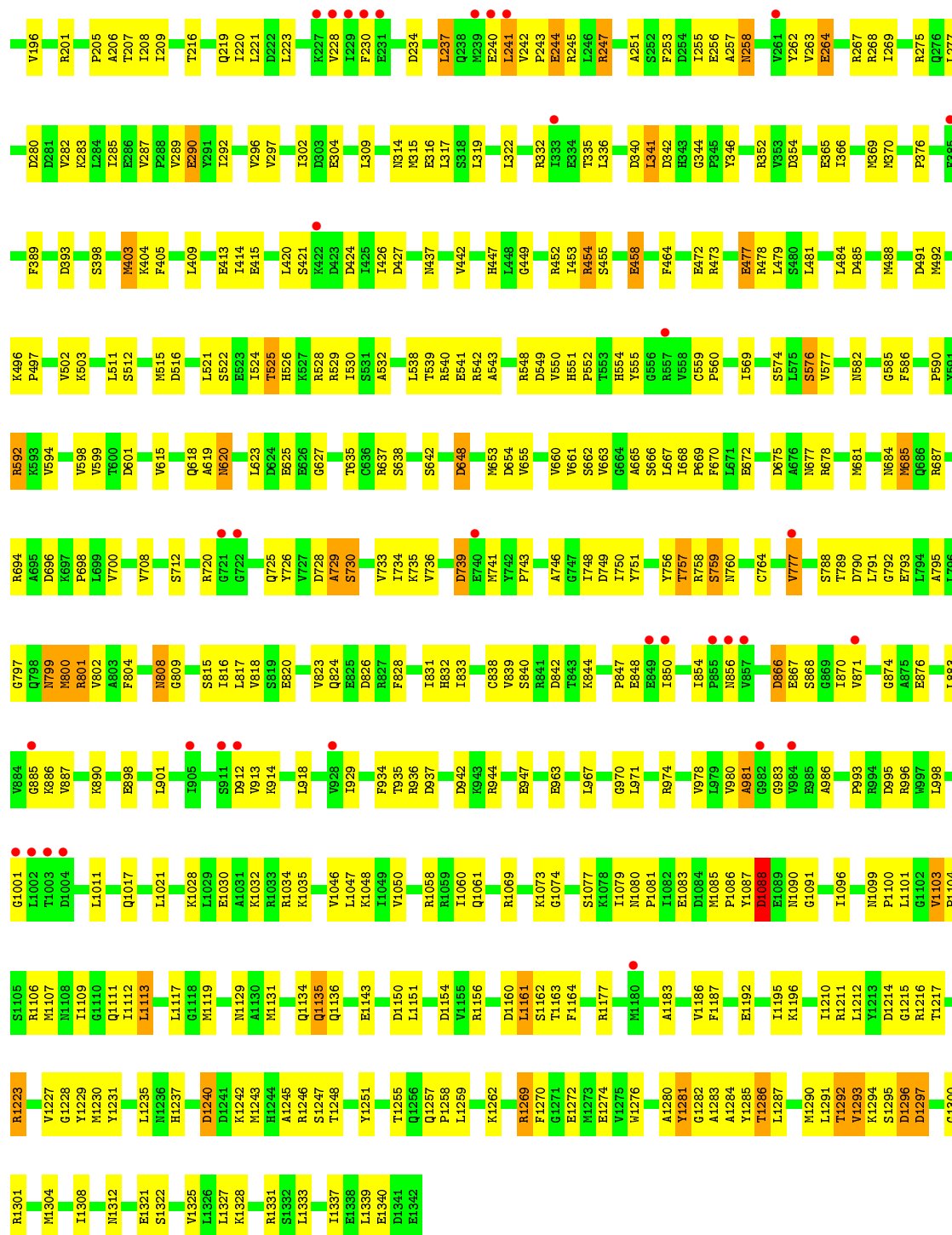


- Molecule 1: DNA-directed RNA polymerase subunit alpha



- Molecule 2: DNA-directed RNA polymerase subunit beta





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



E1343	V1257	V1163	P1063	S977	V885	V808	R731	L614	V518	N424	G336	J0209	V97
L1344	R1288	S1164	S1064	T980	V886	V809	R732	T617	N519	R425	R337	S210	R98
R1345	Q1259	F1165	A1065	E981	S887	T810	S732	V618	A520	A426	F338	E211	R99
K1348	M1260	G1166	E1066	L982	C888	D812	S733	Q736	E523	T428	R339	T212	E100
E1349	K1263	K1167	R1067	L983	D889	T816	Q739	Q736	E524	R430	Q340	R213	R101
H1350	T1266	T1169	D1073	L984	C893	H817	Q739	Q739	N525	H430	Q341	K214	A108
V1351	K1170	K1170	L1078	D986	V894	E818	Q739	Q739	V526	L342	L342	R220	S109
I1352	R1173	R1174	K1079	D986	C895	G819	A741	Q632	T527	G344	G344	L221	P110
I1357	R1174	R1174	L1080	R990	A896	I820	A741	Q632	T528	G345	G345	K222	T111
D1368	L1175	L1175	L1081	T991	H897	N821	R744	Q632	G529	R346	R346	A112	A112
A1369	V1176	V1176	D1082	K992	C898	N822	R744	Q632	P530	V347	V347	L224	H113
A1370	P1179	P1179	A1083	E993	K911	V825	L746	Q632	E532	D348	D348	F227	I114
R1372	Y1186	Y1186	Q1084	S994	S994	A748	R747	A637	G444	T356	T356	M237	K118
R1373	E1187	E1187	D1087	Y995	I915	G829	R749	S638	R535	V357	V357	L245	S119
G1376	E1188	E1188	L1089	P998	I918	D830	P750	L536	Y537	Y360	Y360	P246	L120
GLU	K1192	K1192	L1090	Y999	I918	E833	D751	G640	R538	L361	L361	P247	P121
ALA	W1193	W1193	P1091	G1000	I923	P834	I754	D642	S543	L363	L363	D248	S122
ALA	L1196	L1196	A1097	L1003	G924	R836	P758	V645	V548	Q365	Q365	L249	R123
ALA	W1197	W1197	F1100	A1018	E925	R837	I759	V661	F461	G366	G366	D256	L127
PRO	F1199	F1199	G1103	H1019	P926	D837	F763	I684	D464	G367	G367	M130	M130
GLN	F1199	F1199	G1103	H1019	T928	R838	R764	I684	Q465	R370	R370	A261	D134
VAL	E1200	E1200	K1104	L830	Q929	V839	R764	Q685	M466	K371	K371	T262	I135
THR	G1201	G1201	A1105	P1022	T931	L840	V769	Q686	Q476	N372	N372	R278	E136
ALA	E1202	E1202	L1106	H1023	M932	T844	L770	Q687	L478	E375	E375	L279	L139
GLU	R1203	R1203	E1110	T1024	R933	A845	T776	Q688	L479	L376	L376	L283	L139
ASP	G1207	G1207	D1111	H1025	F935	E846	T776	Q689	L479	F377	F377	S143	S143
ALA	V1208	V1208	V1112	I1028	H936	D847	A779	S670	E475	K378	K378	P288	M151
SER	G1308	G1308	V1113	T1029	L849	V848	R781	V673	L579	P379	P379	D289	E151
SER	I1309	I1309	E1030	E1030	K850	L849	K781	Q685	W580	F380	F380	I290	M151
LEU	T1310	T1310	V1031	S1032	P851	K851	G782	R678	W581	I381	I381	E295	Q157
ALA	S1313	S1313	G1118	S1032	S942	G852	L783	V679	L582	Y382	Y382	Q300	Q158
GLU	L1314	L1314	L1121	F1037	R943	T853	A784	Q690	W583	L385	L385	L160	L160
LEU	S1318	S1318	A1122	T1038	S949	A854	D785	R681	P584	E386	E386	T161	T161
ASN	F1319	F1319	S1128	D1039	S949	D855	T786	I685	G586	L387	L387	V303	Q164
GLY	T1320	T1320	G1129	M1040	Q951	L857	T790	Q686	Q589	L491	L491	D304	Q164
LEU	H1227	H1227	D1042	I1041	V952	V858	A791	M697	S589	S492	S492	A305	D167
GLY	T1230	T1230	L1133	D1042	S957	P859	W792	M698	S590	T499	T499	R311	G173
SER	Q1238	Q1238	T1134	Q1044	I958	R860	G794	Q698	V592	I500	I500	R312	G173
ASP	R1242	R1242	L1138	T1047	K959	T862	W795	Q702	A595	P501	P501	I316	E175
ASN	L1243	L1243	E1146	D1051	V963	L864	T797	E704	S503	S503	S503	I316	F176
GLU	Q1244	Q1244	K1151	E1052	K964	L872	R798	W706	V506	V506	V506	S319	A182
GLY	V1246	V1246	L1156	L1054	S965	E873	R799	T707	L510	E414	E414	N320	A182
ASP	M1249	M1249	A1157	G1055	S969	V877	L800	W708	Y511	V415	V415	K321	I185
ASN	D1250	D1250	E1158	G1055	S970	D878	W801	R709	Y511	E418	E418	P323	E197
GLU	K1340	K1340	G1161	L1059	V974	D878	W801	D710	T514	V421	V421	K332	C198
D1342	V1255	V1255	I1162	V1060	I975	A879	Q805	K715	R515	L422	L422	G333	L205
				V1061	K381	V880	D806	Q716	D516	L422	L422	K334	L205
				L1062	T976	V882	L807	V717	O613	C517	C517	Q335	

• Molecule 4: DNA-directed RNA polymerase subunit omega

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	132.40Å 154.10Å 232.15Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.08 – 4.20 49.08 – 4.20	Depositor EDS
% Data completeness (in resolution range)	98.9 (49.08-4.20) 99.0 (49.08-4.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.08 (at 4.14Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, R_{free}	0.343 , 0.389 0.330 , 0.374	Depositor DCC
R_{free} test set	1678 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å ²)	165.5	Xtriage
Anisotropy	0.130	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.23 , 184.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.29$, $\langle L^2 \rangle = 0.13$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	29026	wwPDB-VP
Average B, all atoms (Å ²)	272.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.42% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: UTP, DOC, MG, 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	AAA	0.62	0/1809	0.76	0/2450
1	BBB	0.62	0/1789	0.74	0/2425
2	CCC	0.62	0/10746	0.82	3/14499 (0.0%)
3	DDD	0.62	0/10729	0.79	0/14487
4	EEE	0.62	0/629	0.81	0/847
5	FFF	0.64	0/2282	0.66	0/3076
6	111	0.27	0/691	0.63	0/1063
7	222	0.32	0/802	0.66	0/1234
8	333	0.16	0/69	0.52	0/106
All	All	0.61	0/29546	0.78	3/40187 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
2	CCC	0	2

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	CCC	1088	ASP	CB-CA-C	-6.16	98.08	110.40
2	CCC	1048	LYS	CB-CA-C	-5.83	98.74	110.40
2	CCC	454	ARG	CB-CA-C	-5.04	100.33	110.40

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	CCC	1282	GLY	Peptide
2	CCC	376	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	AAA	1787	0	1813	85	0
1	BBB	1767	0	1789	100	1
2	CCC	10577	0	10591	459	0
3	DDD	10568	0	10782	508	2
4	EEE	627	0	634	16	0
5	FFF	2253	0	2298	157	0
6	111	618	0	341	44	1
7	222	716	0	397	42	0
8	333	62	0	33	0	0
9	CCC	29	0	11	14	0
10	DDD	2	0	0	2	0
11	DDD	2	0	0	0	0
12	333	18	0	12	7	0
All	All	29026	0	28701	1173	2

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 20.

The worst 5 of 1173 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)
2:CCC:936:ARG:CG	2:CCC:937:ASP:H	1.31	1.35
6:111:54:DA:H2"	6:111:55:DC:C5	1.61	1.33
2:CCC:1106:ARG:NH1	9:CCC:1401:UTP:O3G	1.69	1.22
3:DDD:392:THR:HG21	5:FFF:320:GLN:O	1.37	1.22
1:BBB:83:LEU:HG	3:DDD:526:VAL:CG1	1.69	1.20

All (2) 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)
3:DDD:212:THR:OG1	6:111:27:DC:OP1[3_644]	1.88	0.32
1:BBB:159:ILE:CD1	3:DDD:1054:THR:O[4_445]	2.18	0.02

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	AAA	228/242 (94%)	209 (92%)	14 (6%)	5 (2%)	6	38
1	BBB	226/242 (93%)	209 (92%)	11 (5%)	6 (3%)	5	34
2	CCC	1339/1342 (100%)	1236 (92%)	78 (6%)	25 (2%)	8	41
3	DDD	1360/1407 (97%)	1251 (92%)	90 (7%)	19 (1%)	11	47
4	EEE	77/90 (86%)	73 (95%)	4 (5%)	0	100	100
5	FFF	275/336 (82%)	253 (92%)	16 (6%)	6 (2%)	6	38
All	All	3505/3659 (96%)	3231 (92%)	213 (6%)	61 (2%)	9	44

5 of 61 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	BBB	117	HIS
1	BBB	119	GLY
1	BBB	193	GLU
1	BBB	194	GLN
2	CCC	46	GLN

5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	AAA	198/208 (95%)	178 (90%)	20 (10%)	7	28
1	BBB	196/208 (94%)	180 (92%)	16 (8%)	11	37
2	CCC	1156/1157 (100%)	1066 (92%)	90 (8%)	12	39
3	DDD	1135/1168 (97%)	1063 (94%)	72 (6%)	18	45
4	EEE	67/74 (90%)	63 (94%)	4 (6%)	19	47
5	FFF	240/292 (82%)	228 (95%)	12 (5%)	24	51
All	All	2992/3107 (96%)	2778 (93%)	214 (7%)	14	41

5 of 214 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	CCC	801	ARG
2	CCC	1248	THR
3	DDD	1345	ARG
2	CCC	815	SER
2	CCC	995	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	333	2/4 (50%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 4 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
9	UTP	CCC	1401	11	26,30,30	1.87	4 (15%)	34,47,47	1.01	1 (2%)
12	DOC	333	101	-	14,19,20	0.75	0	13,26,29	1.95	4 (30%)

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	UTP	CCC	1401	11	-	3/22/38/38	0/2/2/2
12	DOC	333	101	-	-	2/4/18/19	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	CCC	1401	UTP	C6-C5	-6.78	1.34	1.52
9	CCC	1401	UTP	C6-N1	-5.20	1.37	1.47
9	CCC	1401	UTP	C5-C4	-2.70	1.43	1.50
9	CCC	1401	UTP	C2-N1	2.07	1.38	1.35

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
12	333	101	DOC	C2-N3-C4	4.31	120.70	116.34
9	CCC	1401	UTP	C5-C6-N1	3.29	122.44	111.61
12	333	101	DOC	N4-C4-N3	2.62	120.63	116.49
12	333	101	DOC	C2'-C1'-N1	-2.41	107.94	112.48
12	333	101	DOC	C5-C4-N3	-2.01	119.40	121.72

There are no chirality outliers.

All (5) torsion outliers are listed below:

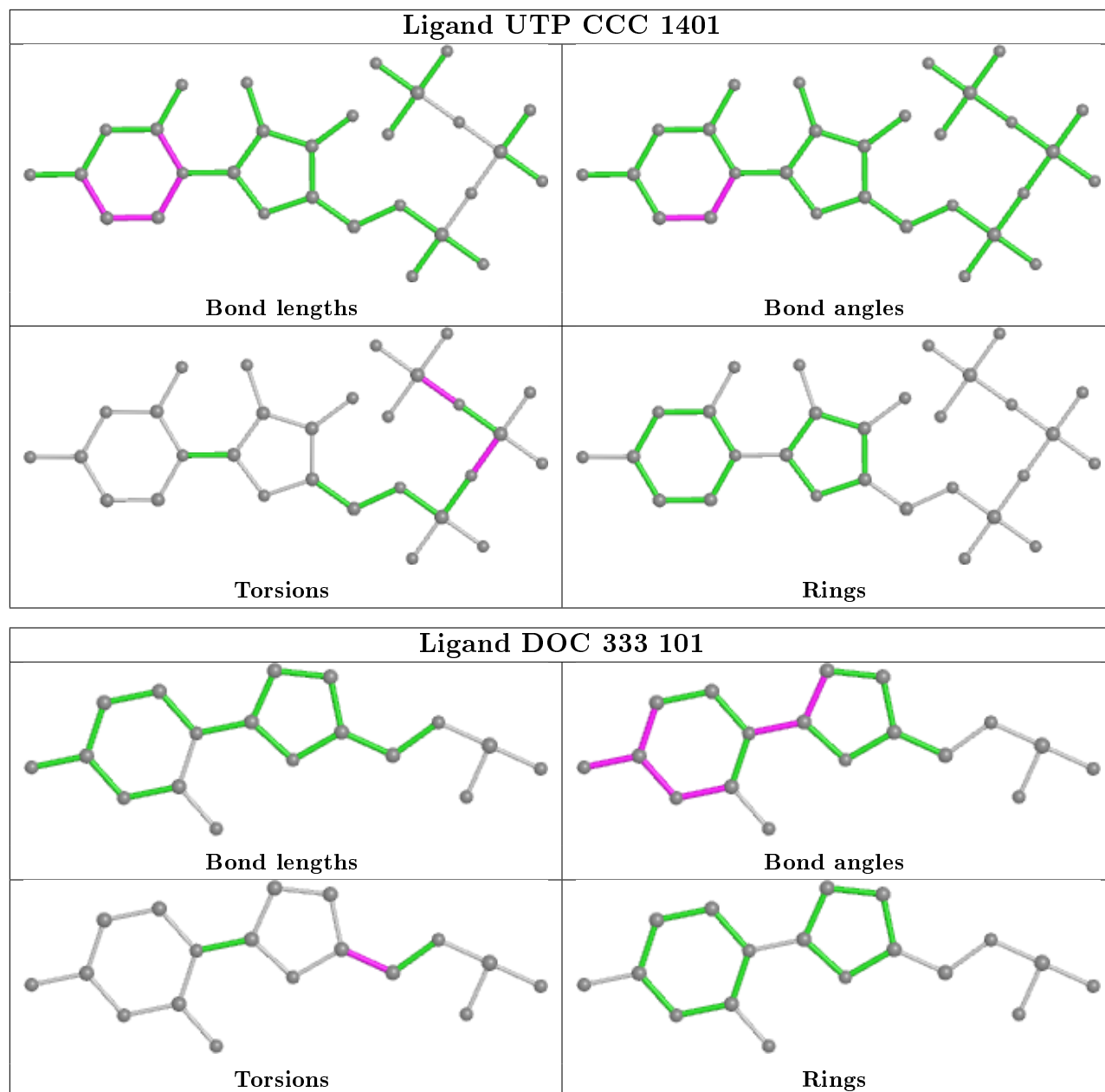
Mol	Chain	Res	Type	Atoms
12	333	101	DOC	C3'-C4'-C5'-O5'
9	CCC	1401	UTP	PB-O3B-PG-O3G
12	333	101	DOC	O4'-C4'-C5'-O5'
9	CCC	1401	UTP	PA-O3A-PB-O1B
9	CCC	1401	UTP	PA-O3A-PB-O2B

There are no ring outliers.

2 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
9	CCC	1401	UTP	14	0
12	333	101	DOC	7	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	AAA	230/242 (95%)	-0.01	8 (3%) 44 35	212, 294, 356, 386	0
1	BBB	228/242 (94%)	0.11	11 (4%) 30 26	215, 288, 363, 434	0
2	CCC	1341/1342 (99%)	-0.14	40 (2%) 50 39	102, 235, 396, 487	0
3	DDD	1362/1407 (96%)	-0.08	51 (3%) 41 33	115, 257, 386, 480	0
4	EEE	79/90 (87%)	-0.05	4 (5%) 28 24	207, 309, 472, 518	0
5	FFF	277/336 (82%)	0.06	7 (2%) 57 47	197, 306, 438, 485	0
6	111	30/50 (60%)	0.06	4 (13%) 3 4	269, 335, 440, 546	0
7	222	35/50 (70%)	0.23	3 (8%) 10 9	267, 350, 509, 513	0
8	333	3/4 (75%)	0.94	0 100 100	351, 351, 399, 414	0
All	All	3585/3763 (95%)	-0.07	128 (3%) 42 34	102, 265, 407, 546	0

The worst 5 of 128 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
6	111	49	DG	6.0
3	DDD	854	ALA	5.9
3	DDD	993	GLU	5.9
2	CCC	1001	GLY	5.2
3	DDD	943	ARG	5.0

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 ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

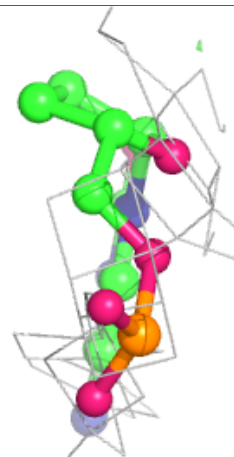
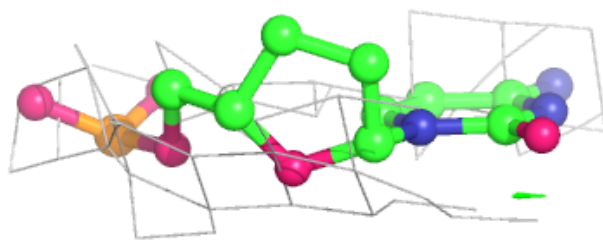
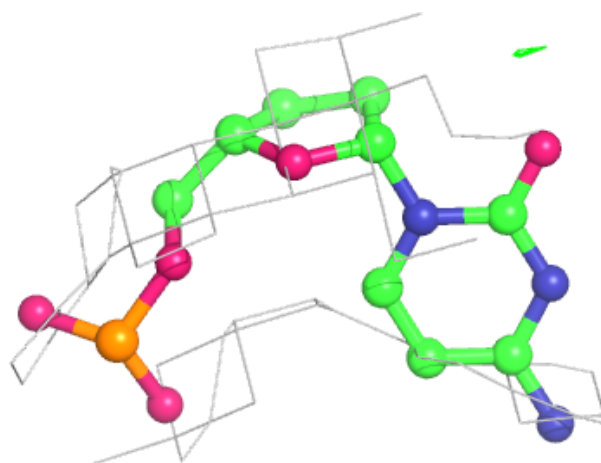
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<0.9' lists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

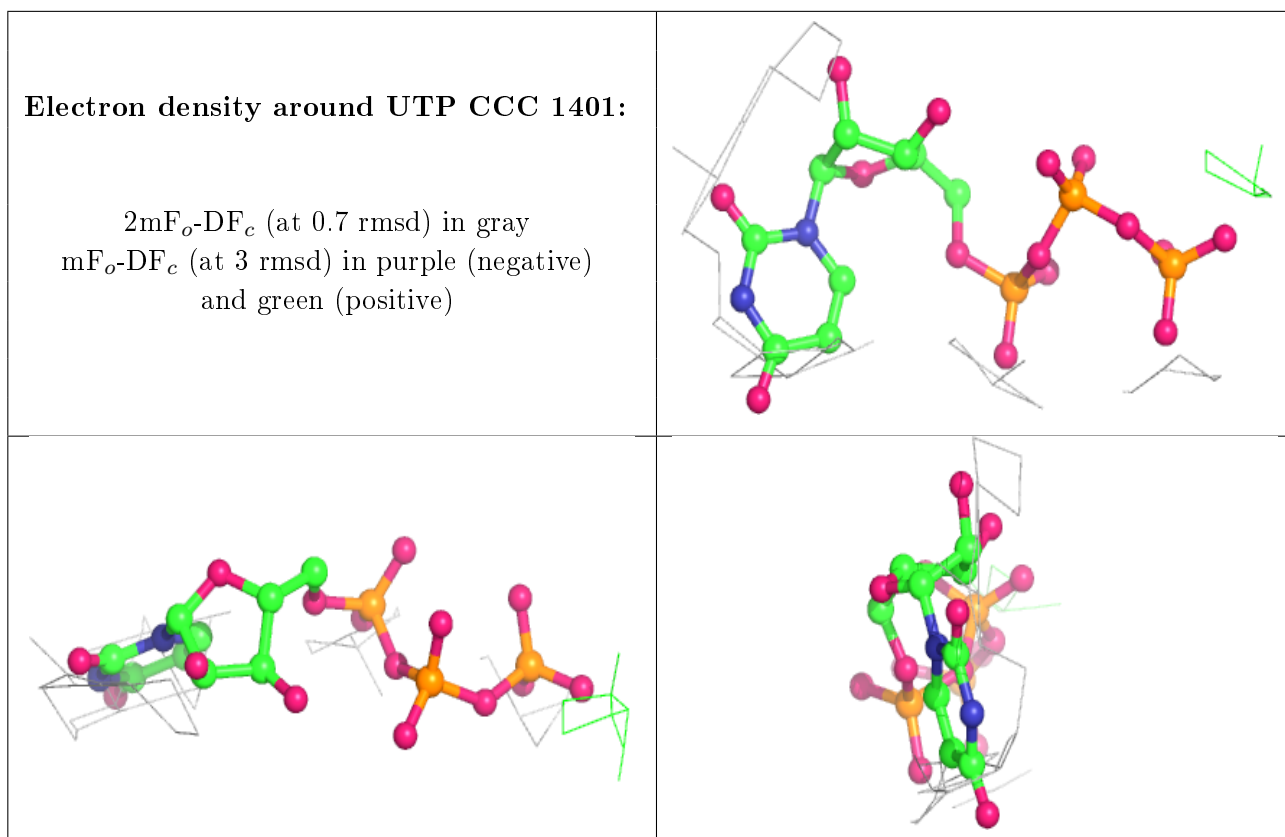
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
12	DOC	333	101	18/19	0.82	0.36	311,333,341,343	0
10	ZN	DDD	1501	1/1	0.88	0.02	423,423,423,423	0
11	MG	DDD	1504	1/1	0.89	0.20	148,148,148,148	0
9	UTP	CCC	1401	29/29	0.90	0.37	239,278,320,322	0
11	MG	DDD	1503	1/1	0.95	0.25	170,170,170,170	0
10	ZN	DDD	1502	1/1	0.99	0.12	194,194,194,194	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 DOC 333 101:

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.