



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 7, 2022 – 04:08 AM EST

PDB ID : 6UU3
Title : E. coli sigma-S transcription initiation complex with a 4-nt RNA and a CTP
("Old" crystal soaked with GTP, ATP, CTP, and ddTTP for 30 minutes)
Authors : Zuo, Y.; De, S.; Steitz, T.A.
Deposited on : 2019-10-30
Resolution : 4.00 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 : 2.27
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) : 2.27

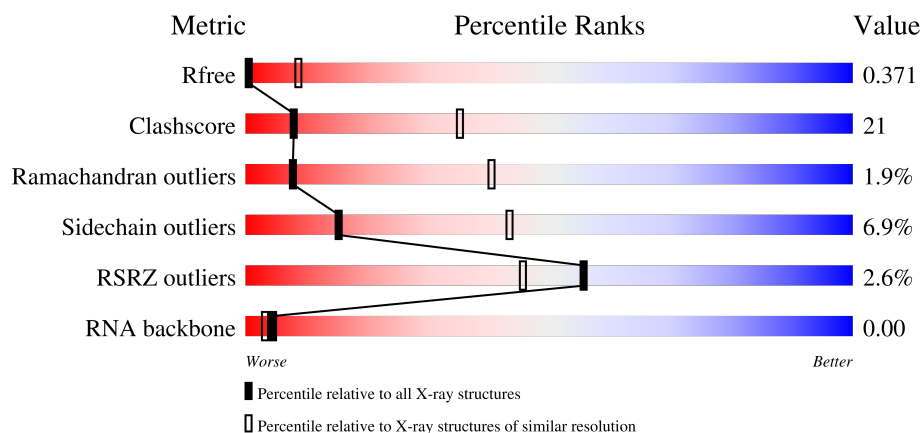
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.00 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	1087 (4.30-3.70)
Clashscore	141614	1148 (4.30-3.70)
Ramachandran outliers	138981	1108 (4.30-3.70)
Sidechain outliers	138945	1099 (4.30-3.70)
RSRZ outliers	127900	1028 (4.34-3.66)
RNA backbone	3102	1048 (5.00-3.00)

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 ≥ 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>2%</div> <div>61% 31% 5%</div> </div>
1	BBB	242	<div> <div>4%</div> <div>62% 29% 6%</div> </div>
2	CCC	1342	<div> <div>%</div> <div>64% 32%</div> </div>
3	DDD	1407	<div> <div>3%</div> <div>58% 36%</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
11	CTP	DDD	1505	-	-	X	-

2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 29064 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 A0A377D9Q8
AAA	-5	HIS	-	expression tag	UNP A0A377D9Q8
AAA	-4	HIS	-	expression tag	UNP A0A377D9Q8
AAA	-3	HIS	-	expression tag	UNP A0A377D9Q8
AAA	-2	HIS	-	expression tag	UNP A0A377D9Q8
AAA	-1	HIS	-	expression tag	UNP A0A377D9Q8
AAA	0	HIS	-	expression tag	UNP A0A377D9Q8
BBB	-6	ALA	-	expression tag	UNP A0A377D9Q8
BBB	-5	HIS	-	expression tag	UNP A0A377D9Q8
BBB	-4	HIS	-	expression tag	UNP A0A377D9Q8
BBB	-3	HIS	-	expression tag	UNP A0A377D9Q8
BBB	-2	HIS	-	expression tag	UNP A0A377D9Q8
BBB	-1	HIS	-	expression tag	UNP A0A377D9Q8
BBB	0	HIS	-	expression tag	UNP A0A377D9Q8

- 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
			10576	6636	1842	2055	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
			619	294	114	181	30			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	222	36	Total	C	N	O	P	0	0	0
			738	352	137	214	35			

- Molecule 8 is a RNA chain called RNA 4-mer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	333	3	Total	C	N	O	P	0	0	0
			77	30	15	27	5			

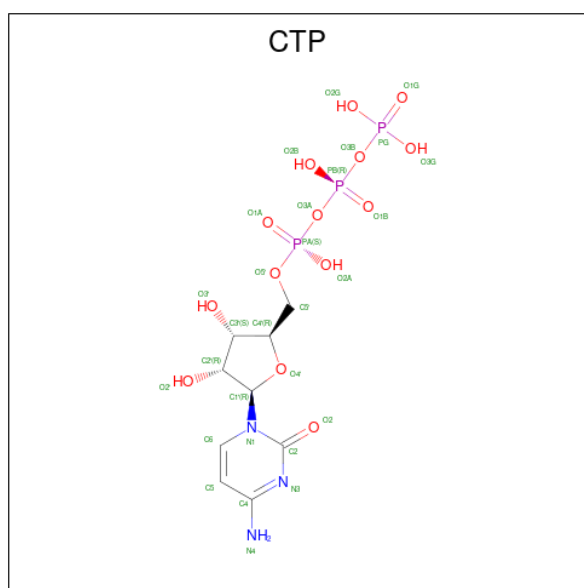
- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

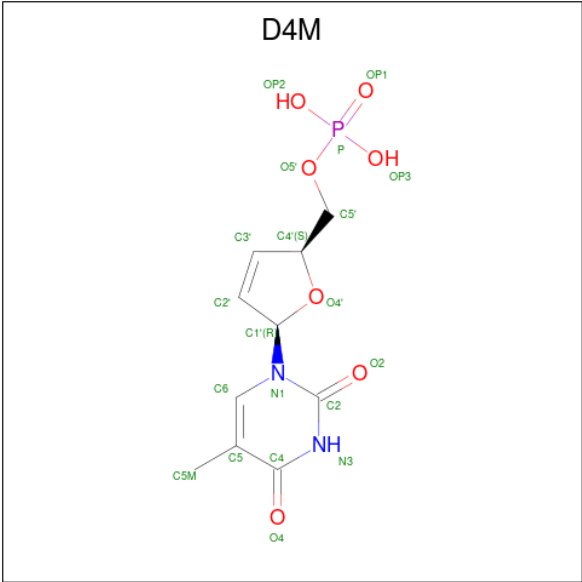
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	DDD	2	Total	Zn	0	0
			2	2		

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

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

- Molecule 11 is CYTIDINE-5'-TRIPHOSPHATE (three-letter code: CTP) (formula: $C_9H_{16}N_3O_{14}P_3$).



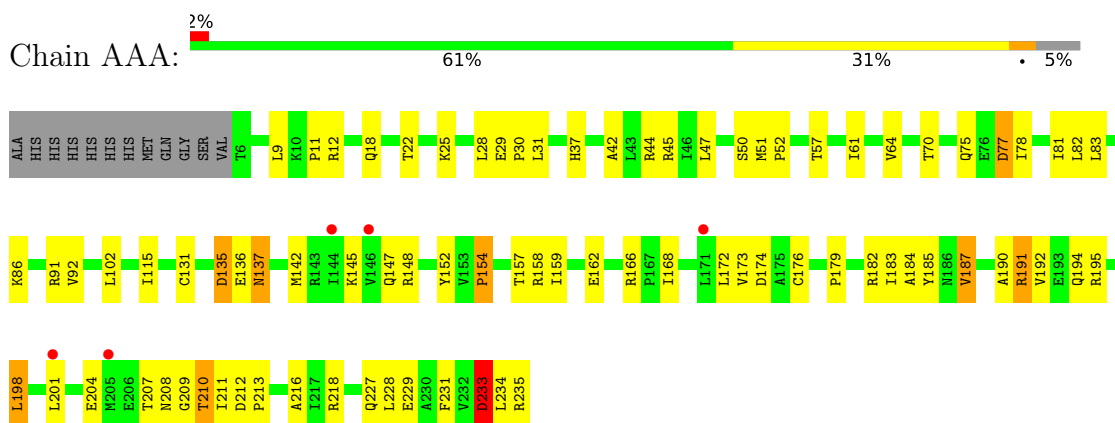


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
12	333	1	Total	C	N	O	P	0	0
			19	10	2	6	1		

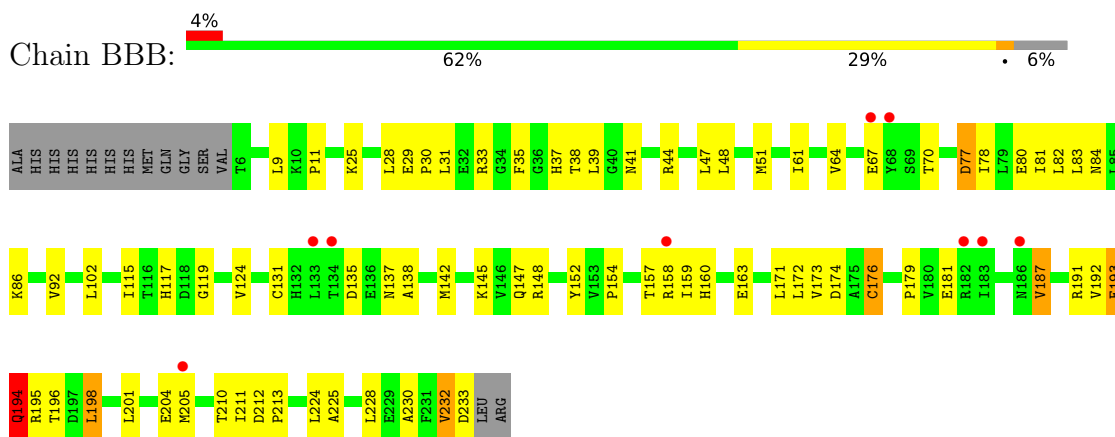
3 Residue-property plots [i](#)

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.

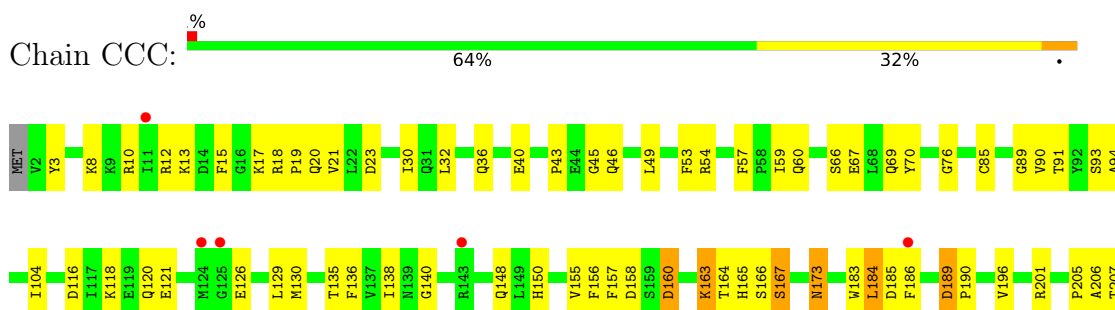
- Molecule 1: DNA-directed RNA polymerase subunit alpha

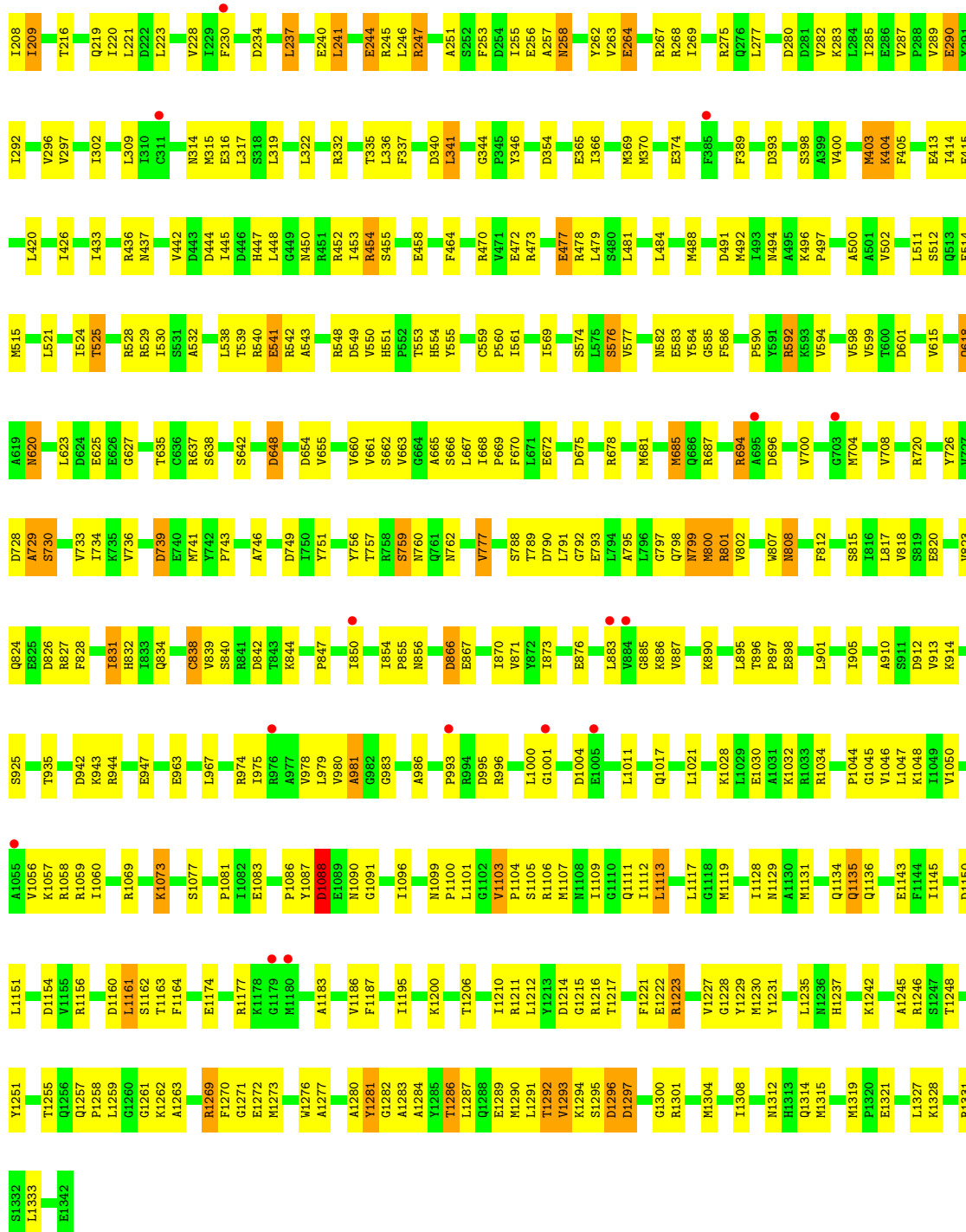


- Molecule 1: DNA-directed RNA polymerase subunit alpha



- Molecule 2: DNA-directed RNA polymerase subunit beta





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



L1307	E1202	L1109	H1023	T934	L857	T786	M697	K598	Y512	L429	G344	A108
L1308	R1203	E1110	T1024	F935	V858	A787	M698	K599	M513	H430	K345	S109
I1309	V1204	D1111	M1025	H936	L788	L788	P859	R860	T514	R431	R346	P110
T1310	E1206	G1112	P1026	I937	R860	K789	L701	S602	D516	I434	V347	T111
K1311	R1205	V1113	G1027	G938	R861	T790	T702	K603	C517	Q435	D348	H112
A1312	G1207	G1118	I1028	G939	T862	A791	T703	M604	V518	R352	P237	H113
S1313	D1208	G1119	E1029	A940	L863	N792	E704	C608	R519	I442	L245	W115
T1316	V1209	L1121	E1030	A941	L864	S793	T705	R610	A520	E443	P246	K118
F1319	I1210	A1122	V1031	S942	H865	T795	I707	R610	G444	P247	D248	S119
S1211	S1211	E1122	S1032	R943	V868	L796	T708	I611	K445	K445	D948	L120
				A944	C869	T797	R709	L612	Q448	L452	D256	S122
				A945	R874	R797	D710	G613	L456	L456	R259	L127
				A946	L872	R798	D710	G613	Y457	Y457	F260	L127
				E947	E873	R799	Q716	M525	A458	A458	A261	D134
				S948	R874	L800	V717	V526	A459	A459	L265	I135
				S949	R875	V801	Q716	L527	D460	D460	K370	E136
				I950	S876	D802	V717	T528	F461	F461	R278	L139
				K955	V876	V803	I722	K531	D462	D462	L279	
				G956	D877	A804	Y723	E532	G463	G463	L283	S143
				S957	D878	Q805	M724	F620	Y457	Y457	A261	
				I958	V880	D806	A621	A533	A458	A458	L265	
				K959	R881	L807	R731	Q623	D622	D622	K370	
				L960	V882	V808	G732	I624	R535	R535	R278	
				V963	R883	T810	S733	M625	L536	D462	L279	
				K964	E884	E811	Q736	T627	D464	D464	L283	
				S965	V885	D812	N1249	G628	Q465	Q465	I290	M151
				S969	R886	L1060	Q739	F629	A467	A467	I291	N153
				S970	C888	V1061	L740	R530	Y468	Y468	E295	
				V974	D889	T816	A741	Y631	L472	L472	Q300	I159
				T975	G893	E819	R744	A632	T473	T473	L160	T161
				S976	R894	I820	G745	R634	L474	L474	L385	
				S977	C895	M822	L746	S635	E475	E475	V303	
				R978	S977	T823	M747	G636	A476	A476	D304	Q164
				L982	C898	P824	P750	V639	Q477	Q477	A305	
				K992	R901	V825	D751	I641	L478	L478	R311	D167
				E993	L903	D830	G752	G640	E479	E479	R312	G173
				S994	D902	V831	S763	K644	T487	T487	I316	E175
				K996	L903	K332	I754	V645	M488	M488	K395	F176
				V997	K911	K332	P758	A577	I489	I489	S319	
				G1000	I915	L835	I759	I664	L490	L490	N320	A182
				G1000	I918	R836	N762	Q665	W580	W580	K321	
				G1000	I918	V839	F763	E666	M581	P493	K325	I185
				G1000	I918	V839	F763	Q667	I582	I411	K325	
				G1000	I918	V839	F763	F668	V583	E414	K325	
				G1000	I918	V839	F763	Q669	V583	E414	K325	
				G1000	I918	V839	F763	S670	P584	V415	M330	E197
				G1000	I918	V839	F763	S670	P584	V415	M330	C198
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V415	M330	
				G1000	I918	V839	F763	S670	P584	V		

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	132.56Å 153.58Å 231.21Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.08 – 4.00 49.03 – 4.00	Depositor EDS
% Data completeness (in resolution range)	98.7 (49.08-4.00) 98.8 (49.03-4.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.07 (at 4.00Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, R_{free}	0.303 , 0.380 0.297 , 0.371	Depositor DCC
R_{free} test set	1924 reflections (4.81%)	wwPDB-VP
Wilson B-factor (Å ²)	175.8	Xtriage
Anisotropy	0.600	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 215.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.32$, $\langle L^2 \rangle = 0.16$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	29064	wwPDB-VP
Average B, all atoms (Å ²)	299.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.77% 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: D4M, CTP, MG, ZN, GTP

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	0.64	0/1809	0.74	0/2450
1	BBB	0.65	0/1789	0.74	0/2425
2	CCC	0.64	0/10745	0.81	4/14499 (0.0%)
3	DDD	0.64	0/10729	0.78	0/14487
4	EEE	0.63	0/629	0.80	0/847
5	FFF	0.65	0/2282	0.67	0/3076
6	111	0.31	0/693	0.63	0/1066
7	222	0.33	1/828 (0.1%)	0.60	0/1277
8	333	0.21	0/50	0.60	0/76
All	All	0.63	1/29554 (0.0%)	0.77	4/40203 (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	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	222	26	DT	O3'-P	-5.06	1.55	1.61

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	CCC	1088	ASP	CB-CA-C	-6.29	97.81	110.40
2	CCC	1048	LYS	CB-CA-C	-5.89	98.62	110.40
2	CCC	454	ARG	CB-CA-C	-5.51	99.39	110.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	CCC	1222	GLU	CB-CA-C	5.31	121.02	110.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	CCC	1282	GLY	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	AAA	1787	0	1813	77	0
1	BBB	1767	0	1789	85	0
2	CCC	10576	0	10591	471	0
3	DDD	10568	0	10781	584	0
4	EEE	627	0	634	20	0
5	FFF	2253	0	2298	115	1
6	111	619	0	340	39	0
7	222	738	0	407	65	0
8	333	77	0	33	7	0
9	DDD	2	0	0	1	0
10	DDD	2	0	0	0	0
11	DDD	29	0	12	9	0
12	333	19	0	11	6	0
All	All	29064	0	28709	1237	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 21.

The worst 5 of 1237 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:896:THR:OG1	2:CCC:897:PRO:HD2	1.19	1.34
2:CCC:444:ASP:O	2:CCC:450:ASN:ND2	1.68	1.26

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
8:333:16:G:O3'	12:333:101:D4M:P	1.99	1.20
5:FFF:183:ARG:NH2	7:222:26:DT:OP2	1.76	1.19
6:111:53:DG:H2''	6:111:54:DA:OP2	1.39	1.17

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)
5:FFF:67:TYR:O	5:FFF:299:ARG:NH2[3_644]	1.96	0.24

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	
1	AAA	228/242 (94%)	213 (93%)	10 (4%)	5 (2%)	6	37
1	BBB	226/242 (93%)	207 (92%)	14 (6%)	5 (2%)	6	37
2	CCC	1339/1342 (100%)	1240 (93%)	72 (5%)	27 (2%)	7	40
3	DDD	1360/1407 (97%)	1248 (92%)	90 (7%)	22 (2%)	9	44
4	EEE	77/90 (86%)	73 (95%)	3 (4%)	1 (1%)	12	48
5	FFF	275/336 (82%)	256 (93%)	14 (5%)	5 (2%)	8	41
All	All	3505/3659 (96%)	3237 (92%)	203 (6%)	65 (2%)	8	40

5 of 65 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%)	180 (91%)	18 (9%)	9	33
1	BBB	196/208 (94%)	184 (94%)	12 (6%)	18	47
2	CCC	1156/1157 (100%)	1065 (92%)	91 (8%)	12	39
3	DDD	1135/1168 (97%)	1067 (94%)	68 (6%)	19	47
4	EEE	67/74 (90%)	63 (94%)	4 (6%)	19	47
5	FFF	240/292 (82%)	227 (95%)	13 (5%)	22	50
All	All	2992/3107 (96%)	2786 (93%)	206 (7%)	15	43

5 of 206 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	CCC	1269	ARG
3	DDD	545	HIS
5	FFF	229	ASP
2	CCC	1296	ASP
3	DDD	211	GLU

Sometimes 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	1/4 (25%)	1 (100%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
8	333	16	G

There are no RNA pucker outliers to report.

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

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

5.5 Carbohydrates ⓘ

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	$\# Z > 2$
12	D4M	333	101	-	14,20,21	2.30	3 (21%)	14,28,31	2.80	6 (42%)
11	CTP	DDD	1505	10	23,30,30	0.71	0	30,47,47	1.04	1 (3%)

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
12	D4M	333	101	-	-	0/4/18/19	0/2/2/2
11	CTP	DDD	1505	10	-	5/20/38/38	0/2/2/2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
12	333	101	D4M	C2'-C3'	6.66	1.51	1.32
12	333	101	D4M	C5-C4	3.47	1.48	1.41
12	333	101	D4M	C1'-C2'	2.19	1.52	1.49

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
12	333	101	D4M	C2-N3-C4	6.98	121.04	115.14
12	333	101	D4M	O4'-C4'-C3'	3.83	108.44	103.06
11	DDD	1505	CTP	C4-N3-C2	3.75	120.14	116.34
12	333	101	D4M	O4'-C1'-C2'	3.55	108.86	104.91
12	333	101	D4M	C1'-C2'-C3'	-3.21	105.78	109.68

There are no chirality outliers.

All (5) torsion outliers are listed below:

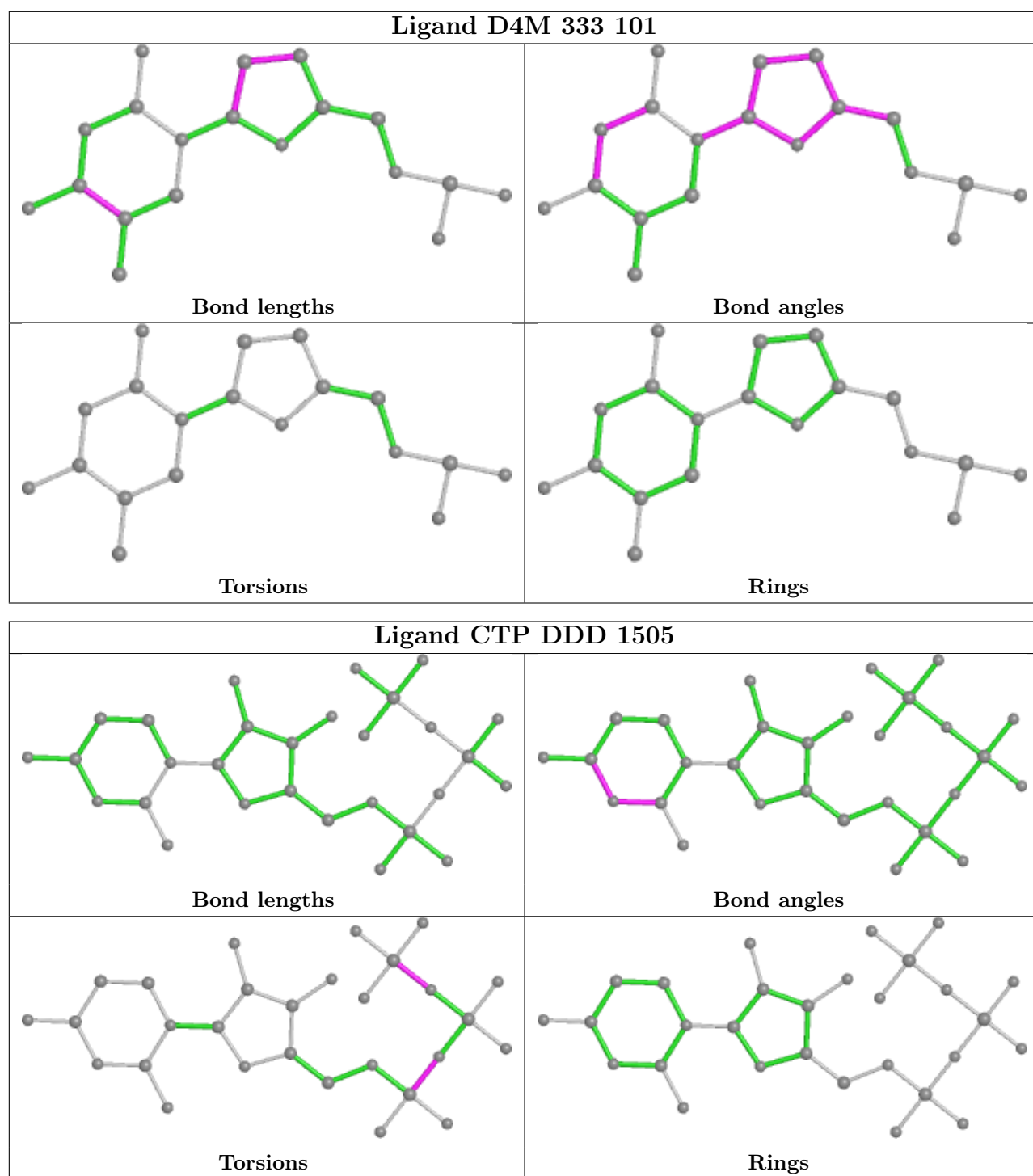
Mol	Chain	Res	Type	Atoms
11	DDD	1505	CTP	PB-O3A-PA-O2A
11	DDD	1505	CTP	PB-O3B-PG-O1G
11	DDD	1505	CTP	PB-O3A-PA-O1A
11	DDD	1505	CTP	PB-O3B-PG-O2G
11	DDD	1505	CTP	PB-O3B-PG-O3G

There are no ring outliers.

2 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
12	333	101	D4M	6	0
11	DDD	1505	CTP	9	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 ⓘ

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, 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%)	-0.03	5 (2%) 62 52	240, 309, 365, 405	0
1	BBB	228/242 (94%)	0.03	9 (3%) 39 31	219, 314, 391, 434	0
2	CCC	1341/1342 (99%)	-0.13	20 (1%) 73 64	157, 272, 392, 510	0
3	DDD	1362/1407 (96%)	-0.06	38 (2%) 53 42	138, 292, 385, 461	0
4	EEE	79/90 (87%)	-0.25	3 (3%) 40 32	265, 333, 444, 488	0
5	FFF	277/336 (82%)	0.14	16 (5%) 23 19	237, 324, 438, 516	0
6	111	30/50 (60%)	0.06	2 (6%) 17 14	300, 335, 406, 466	0
7	222	36/50 (72%)	0.32	2 (5%) 24 21	226, 346, 465, 492	0
8	333	2/4 (50%)	0.52	0 100 100	296, 296, 296, 323	0
All	All	3585/3763 (95%)	-0.06	95 (2%) 56 46	138, 296, 401, 516	0

The worst 5 of 95 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	FFF	306	GLN	6.6
5	FFF	310	LEU	6.4
7	222	38	DG	5.8
6	111	48	DA	5.5
5	FFF	263	LEU	5.3

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](#)

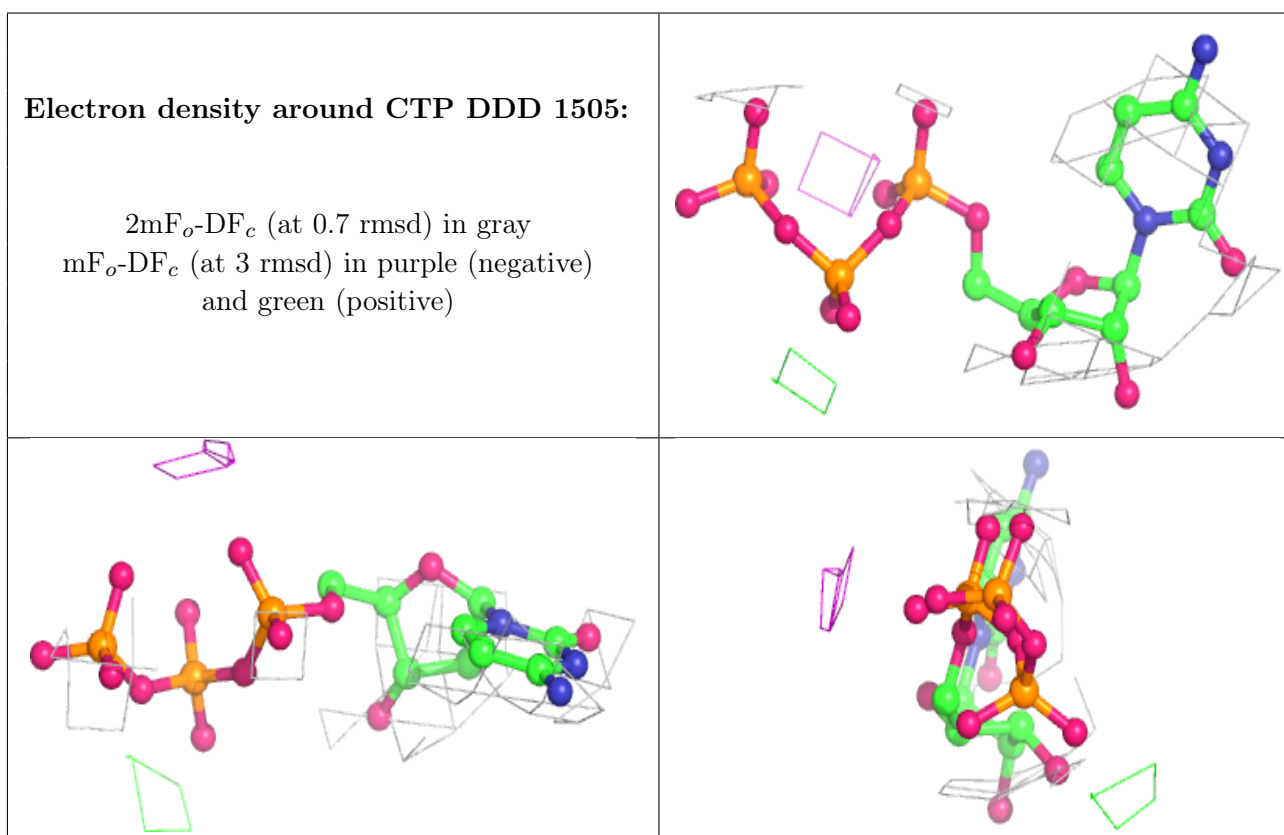
There are no monosaccharides in this entry.

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.

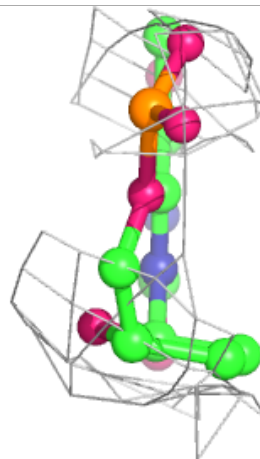
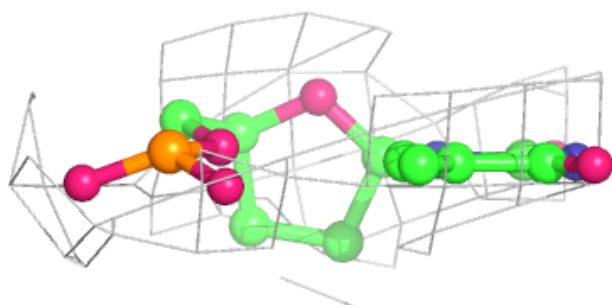
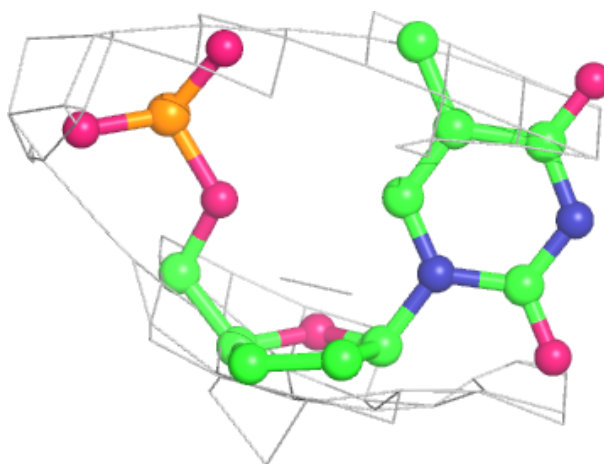
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
11	CTP	DDD	1505	29/29	0.89	0.24	194,235,295,309	0
12	D4M	333	101	19/20	0.93	0.21	242,269,300,304	0
9	ZN	DDD	1501	1/1	0.94	0.10	349,349,349,349	0
10	MG	DDD	1503	1/1	0.94	0.19	108,108,108,108	0
10	MG	DDD	1504	1/1	0.97	0.21	345,345,345,345	0
9	ZN	DDD	1502	1/1	0.98	0.19	337,337,337,337	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 D4M 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.