



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 02:09 PM UTC

PDB ID : 6UU8 / pdb_00006uu8
Title : E. coli mutant sigma-S transcription initiation complex with a 7-nt RNA
("Fresh" mutant crystal soaked with GTP, UTP, and CTP for 30 minutes)
Authors : Zuo, Y.; De, S.; Steitz, T.A.
Deposited on : 2019-10-30
Resolution : 4.40 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
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

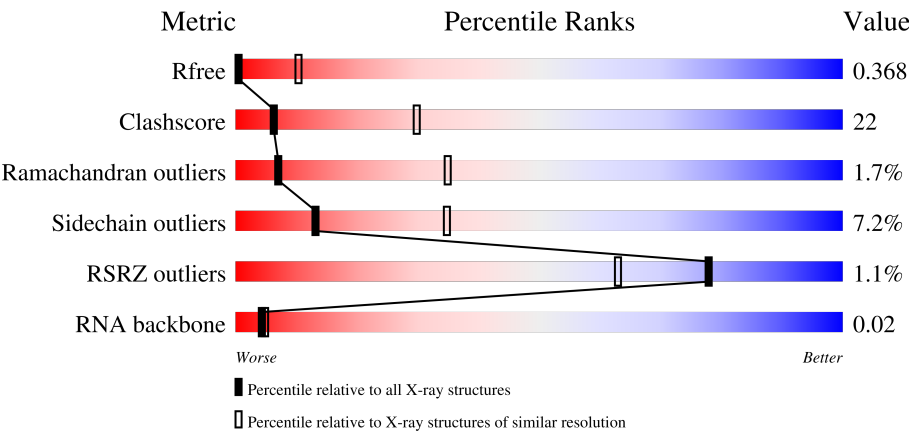
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.40 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	1088 (4.90-3.90)
Clashscore	190562	1135 (4.90-3.90)
Ramachandran outliers	187476	1026 (4.90-3.90)
Sidechain outliers	187428	1010 (4.90-3.90)
RSRZ outliers	180081	1084 (4.90-3.90)
RNA backbone	3983	1037 (5.60-3.10)

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	2% 57%33%5%
1	BBB	242	% 61%29%6%
2	CCC	1342	% 62%33%5%
3	DDD	1407	% 57%35%. .

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	7	

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	DPO	DDD	1504	-	-	X	-

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 28931 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
			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	268	Total	C	N	O	S	0	0	0
			2186	1370	405	407	4			

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
FFF	2	GLY	SER	conflict	UNP A0A377K1M2
FFF	219	GLY	ILE	engineered mutation	UNP A0A377K1M2
FFF	221	ALA	SER	engineered mutation	UNP A0A377K1M2
FFF	329	LEU	-	expression tag	UNP A0A377K1M2
FFF	330	GLU	-	expression tag	UNP A0A377K1M2
FFF	331	HIS	-	expression tag	UNP A0A377K1M2
FFF	332	HIS	-	expression tag	UNP A0A377K1M2
FFF	333	HIS	-	expression tag	UNP A0A377K1M2
FFF	334	HIS	-	expression tag	UNP A0A377K1M2
FFF	335	HIS	-	expression tag	UNP A0A377K1M2
FFF	336	HIS	-	expression tag	UNP A0A377K1M2

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	111	29	Total	C	N	O	P	0	0	0
			595	283	107	176	29			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	222	32	Total	C	N	O	P	0	0	0
			652	312	117	192	31			

- Molecule 8 is a RNA chain called RNA 7-mer (de novo synthesized).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	333	7	Total	C	N	O	P	0	0	0
			160	67	27	57	9			

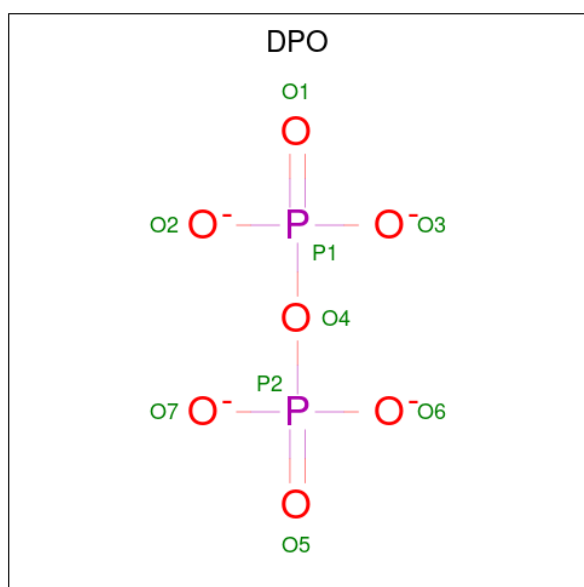
- Molecule 9 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	CCC	1	Total	Mg	0	0
			1	1		
9	DDD	1	Total	Mg	0	0
			1	1		

- Molecule 10 is ZINC ION (CCD ID: ZN) (formula: Zn).

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

- Molecule 11 is DIPHOSPHATE (CCD ID: DPO) (formula: O₇P₂).

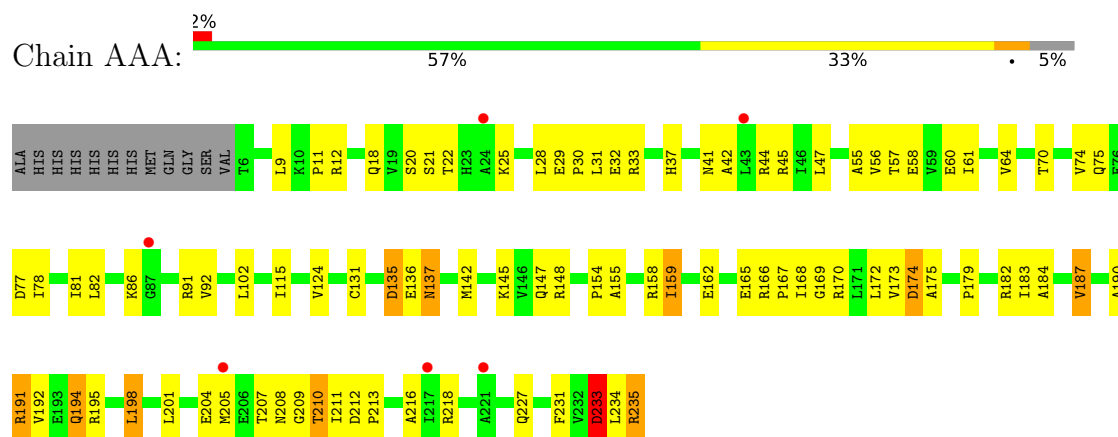


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	DDD	1	Total	O	P	0	0
			9	7	2		

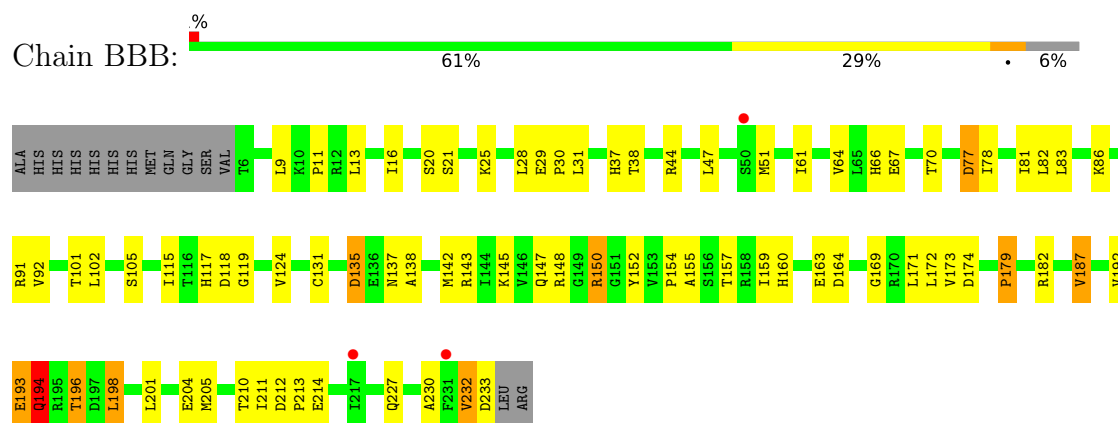
3 Residue-property plots

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.

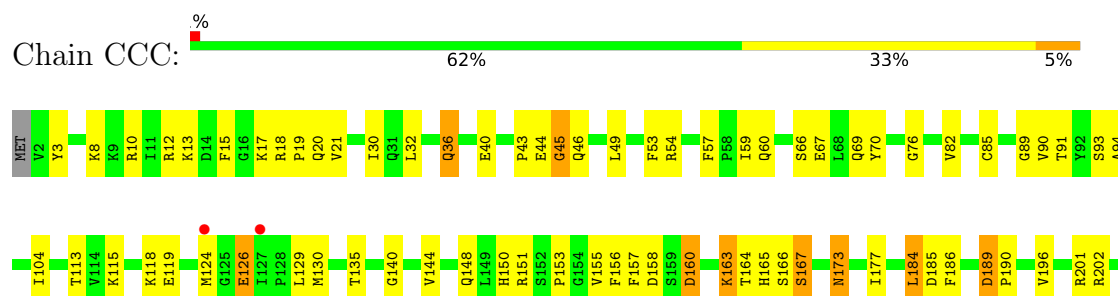
- Molecule 1: DNA-directed RNA polymerase subunit alpha

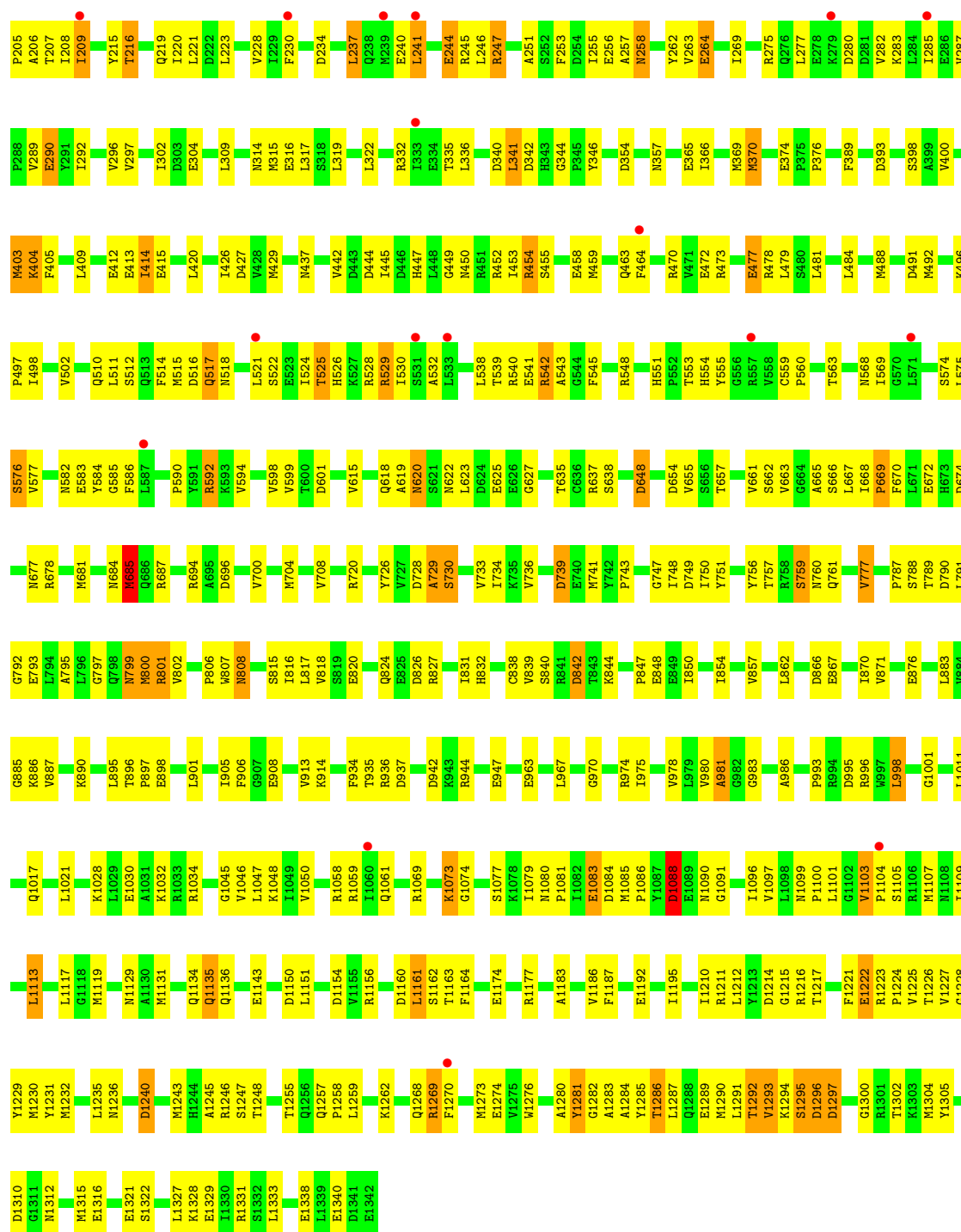


- Molecule 1: DNA-directed RNA polymerase subunit alpha

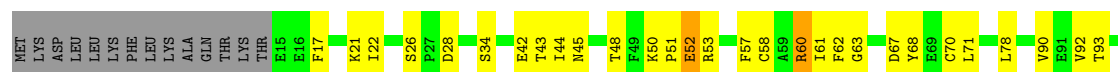


- Molecule 2: DNA-directed RNA polymerase subunit beta





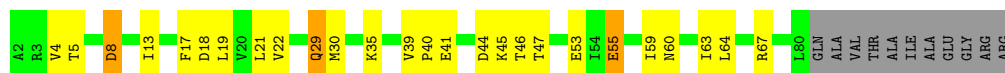
• Molecule 3: DNA-directed RNA polymerase subunit beta'



ASN	ALA	GLY	LEU	GLY	GLY	SER	ASP	ASN	GLU	S994	Y995	K996	Y997	G1000	L1003	A1018	N1019	V1020	D1021	P1022	H1023	T1024	M1025	I1028	T1029	E1030	V1031	S1032	F1037	T1038	D1039	M1040	T1041	H1042	G1043	Q1044	T1047	D1051	E1052	L1053	L1059	V1060	V1061	L1062	D1063	A1065	E1066	R1067	D1073	L1078	K1079	Y1082	E1083	A1084	Q1084	D1087	V1088
L1089	L1090	P1091	G1103	K1104	L1105	A1106	E1110	D1111	G1112	V1113	L1121	S1128	D1133	I1134	T1135	L1138	L1144	F1145	P1150	K1151	E1152	P1153	L1156	A1157	E1158	G1161	L1162	V1163	F1165	E1168	T1169	K1170	R1173	L1174	L1175	V1176	P1179	Y1186	E1187	A1188	W1193	R1194	Q1195														
L1196	L1307	G1308	L1309	L1310	K1311	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU	LEU								
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327	T1328	R1329	R1330	E1334	A1338	G1339	K1340	R1345	V1351	T1357	T1361	D1368	R1369	M1370	R1371	Q1372	R1373	G1376	GLU	ALA	PRO	PRO	ALA	ALA	PRO	GLN	THR	ALA	GLU	ASP	ALA	SER	ALA	SER	LEU	ALA	GLU	LEU									
L1307	L1308	L1309	L1310	L1311	L1312	A1312	S1313	F1319	L1320	E1327																																															

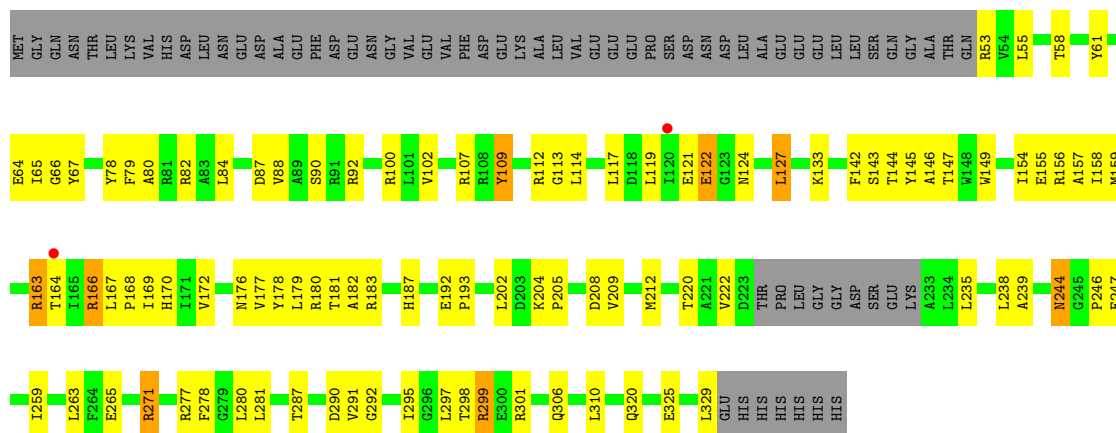
- Molecule 4: DNA-directed RNA polymerase subunit omega

Chain EEE:  59% 26% 12%

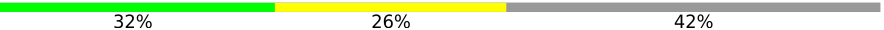


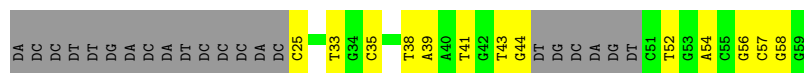
- Molecule 5: RNA polymerase sigma factor RpoS

Chain FFF:  50% 27% 20%



- Molecule 6: Synthetic DNA 50-mer (promoter non-template strand)

Chain 111:  32% 26% 42%

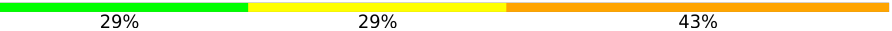


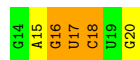
- Molecule 7: Synthetic DNA 50-mer (promoter template strand)

Chain 222:  26% 38% 36%



- Molecule 8: RNA 7-mer (de novo synthesized)

Chain 333:  29% 29% 43%



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	133.46Å 154.50Å 235.70Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.38 – 4.40 49.38 – 4.40	Depositor EDS
% Data completeness (in resolution range)	99.0 (49.38-4.40) 99.0 (49.38-4.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.07 (at 4.45Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, R_{free}	0.357 , 0.380 0.347 , 0.368	Depositor DCC
R_{free} test set	1496 reflections (4.73%)	wwPDB-VP
Wilson B-factor (Å ²)	163.9	Xtriage
Anisotropy	0.949	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 289.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.30$, $\langle L^2 \rangle = 0.14$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	28931	wwPDB-VP
Average B, all atoms (Å ²)	293.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.27% 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: DPO, 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	0.97	0/1809	1.25	2/2450 (0.1%)
1	BBB	1.01	1/1789 (0.1%)	1.29	5/2425 (0.2%)
2	CCC	0.96	0/10745	1.35	32/14499 (0.2%)
3	DDD	0.98	1/10729 (0.0%)	1.33	22/14487 (0.2%)
4	EEE	0.97	0/629	1.45	1/847 (0.1%)
5	FFF	0.98	0/2213	1.41	1/2981 (0.0%)
6	111	0.31	0/665	0.58	0/1022
7	222	0.31	0/729	0.52	0/1121
8	333	0.41	0/142	0.81	0/219
All	All	0.95	2/29450 (0.0%)	1.31	63/40051 (0.2%)

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

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	BBB	77	ASP	C-O	5.73	1.31	1.23
3	DDD	379	PRO	C-O	-5.36	1.17	1.24

The worst 5 of 63 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	CCC	563	THR	CB-CA-C	9.68	123.85	109.63
3	DDD	646	ILE	CA-C-N	9.60	129.68	119.89

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	DDD	646	ILE	C-N-CA	9.60	129.68	119.89
2	CCC	1084	ASP	CA-CB-CG	9.12	121.72	112.60
2	CCC	153	PRO	N-CD-CG	-8.32	90.73	103.20

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	91	0
1	BBB	1767	0	1789	99	0
2	CCC	10576	0	10591	462	0
3	DDD	10568	0	10781	613	0
4	EEE	627	0	634	21	0
5	FFF	2186	0	2230	123	1
6	111	595	0	327	22	0
7	222	652	0	364	39	0
8	333	160	0	75	15	0
9	CCC	1	0	0	0	0
9	DDD	1	0	0	0	0
10	DDD	2	0	0	1	0
11	DDD	9	0	0	2	0
All	All	28931	0	28604	1274	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 22.

The worst 5 of 1274 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)
3:DDD:846:GLU:HG3	3:DDD:860:ARG:NH2	1.42	1.35
2:CCC:896:THR:OG1	2:CCC:897:PRO:HD2	1.22	1.34
3:DDD:525:MET:N	3:DDD:548:VAL:HG23	1.42	1.31
2:CCC:157:PHE:O	2:CCC:442:VAL:HG13	1.31	1.28
3:DDD:846:GLU:CG	3:DDD:860:ARG:NH2	1.96	1.28

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:NH1[3_644]	2.16	0.04

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%)	211 (92%)	11 (5%)	6 (3%)	4	25
1	BBB	226/242 (93%)	206 (91%)	12 (5%)	8 (4%)	3	20
2	CCC	1339/1342 (100%)	1240 (93%)	75 (6%)	24 (2%)	6	33
3	DDD	1360/1407 (97%)	1251 (92%)	91 (7%)	18 (1%)	9	41
4	EEE	77/90 (86%)	73 (95%)	4 (5%)	0	100	100
5	FFF	264/336 (79%)	248 (94%)	14 (5%)	2 (1%)	16	52
All	All	3494/3659 (96%)	3229 (92%)	207 (6%)	58 (2%)	7	35

5 of 58 Ramachandran outliers are listed below:

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

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	CCC	247	ARG

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	28
1	BBB	196/208 (94%)	185 (94%)	11 (6%)	19	41
2	CCC	1156/1157 (100%)	1065 (92%)	91 (8%)	11	32
3	DDD	1135/1168 (97%)	1054 (93%)	81 (7%)	13	35
4	EEE	67/74 (90%)	63 (94%)	4 (6%)	17	39
5	FFF	231/290 (80%)	220 (95%)	11 (5%)	23	45
All	All	2983/3105 (96%)	2767 (93%)	216 (7%)	13	35

5 of 216 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	CCC	1269	ARG
3	DDD	479	GLU
3	DDD	1330	ARG
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	5/7 (71%)	3 (60%)	0

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
8	333	16	G
8	333	17	U
8	333	18	C

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 4 are monoatomic - leaving 1 for Mogul analysis.

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

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
11	DPO	DDD	1504	-	6,8,8	0.74	0	12,13,13	0.81	0

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
11	DPO	DDD	1504	-	-	0/6/6/6	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	DDD	1504	DPO	2	0

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, 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%)	-0.11	6 (2%) 57 44	232, 313, 353, 393	0
1	BBB	228/242 (94%)	-0.14	3 (1%) 75 60	221, 286, 358, 431	0
2	CCC	1341/1342 (99%)	-0.24	19 (1%) 73 58	135, 251, 375, 471	0
3	DDD	1362/1407 (96%)	-0.27	8 (0%) 85 73	151, 282, 429, 476	0
4	EEE	79/90 (87%)	-0.33	0 100 100	238, 294, 444, 493	0
5	FFF	268/336 (79%)	-0.20	2 (0%) 84 71	273, 363, 422, 452	0
6	111	29/50 (58%)	-0.07	0 100 100	281, 404, 468, 485	0
7	222	32/50 (64%)	0.15	0 100 100	291, 388, 452, 489	0
8	333	6/7 (85%)	-0.14	0 100 100	311, 325, 352, 367	0
All	All	3575/3766 (94%)	-0.23	38 (1%) 78 63	135, 281, 419, 493	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	217	ILE	5.3
2	CCC	127	ILE	4.6
2	CCC	333	ILE	4.4
2	CCC	533	LEU	4.4
1	BBB	231	PHE	3.7

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, 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	MG	DDD	1503	1/1	0.14	0.32	192,192,192,192	0
11	DPO	DDD	1504	9/9	0.79	0.10	197,205,216,219	0
9	MG	CCC	1401	1/1	0.93	0.12	143,143,143,143	0
10	ZN	DDD	1502	1/1	0.97	0.07	269,269,269,269	0
10	ZN	DDD	1501	1/1	0.99	0.03	304,304,304,304	0

6.5 Other polymers [i](#)

There are no such residues in this entry.