



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 27, 2020 – 02:06 PM BST

PDB ID : 6UU5  
Title : E. coli sigma-S transcription initiation complex with a 6-nt RNA ("Old" crystal soaked with GTP, UTP, CTP, and dinucleotide GpA for 30 minutes)  
Authors : Zuo, Y.; De, S.; Steitz, T.A.  
Deposited on : 2019-10-30  
Resolution : 5.40 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](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 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
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

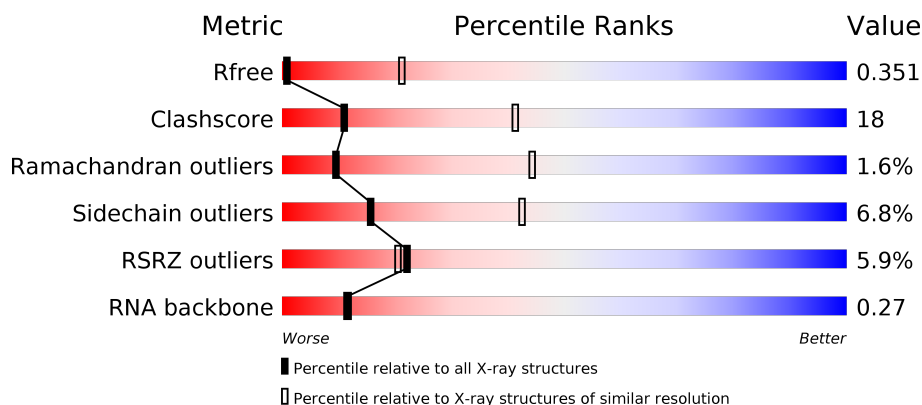
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 5.40 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))
$R_{free}$	130704	1207 (7.00-3.80)
Clashscore	141614	1016 (6.92-3.86)
Ramachandran outliers	138981	1210 (7.00-3.80)
Sidechain outliers	138945	1181 (7.00-3.80)
RSRZ outliers	127900	1021 (7.04-3.76)
RNA backbone	3102	1074 (7.80-3.00)

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  $\geq 3$ , 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	AAA	242	
1	BBB	242	
2	CCC	1342	
3	DDD	1407	

*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	6	

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	-
9	ZN	DDD	1502	-	-	X	-

## 2 Entry composition [i](#)

There are 11 unique types of molecules in this entry. The entry contains 29050 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	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 6-mer (dinucleotide GpA primed synthesis).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	333	6	Total	C	N	O	P	0	0	0
			125	57	22	41	5			

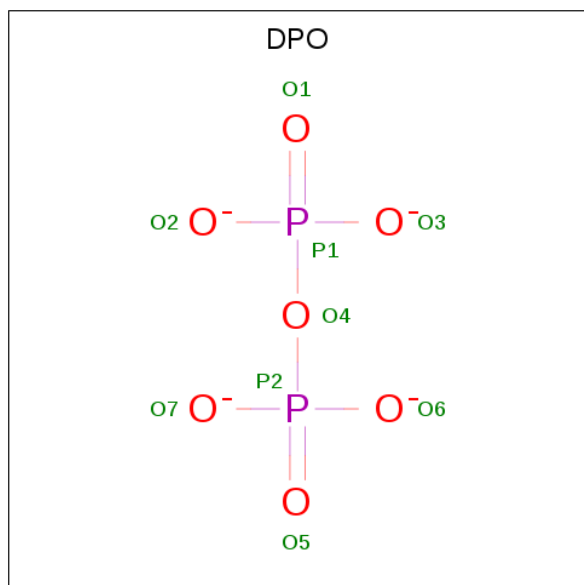
- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

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	333	1	Total	Mg	0	0
			1	1		
10	DDD	1	Total	Mg	0	0
			1	1		

- Molecule 11 is DIPHOSPHATE (three-letter code: DPO) (formula: O<sub>7</sub>P<sub>2</sub>).

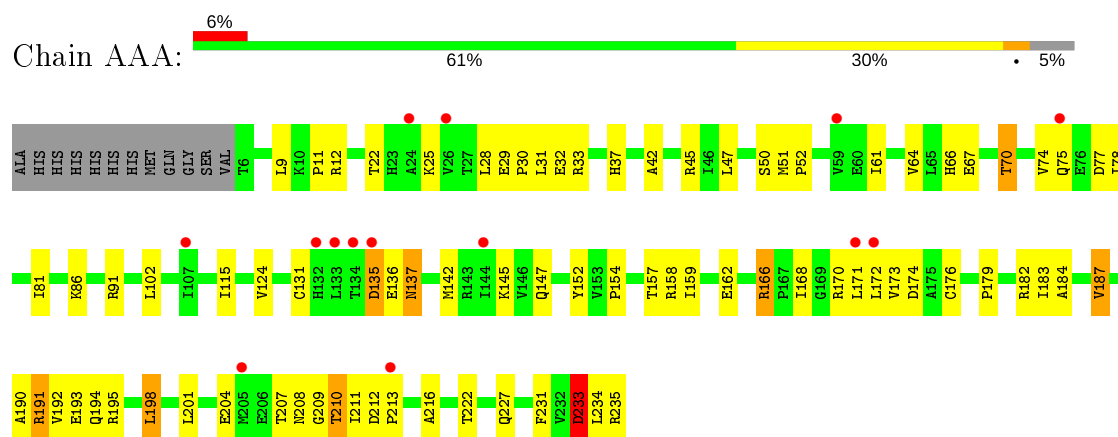


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
11	DDD	1	Total	O	P	0	0
			9	7	2		

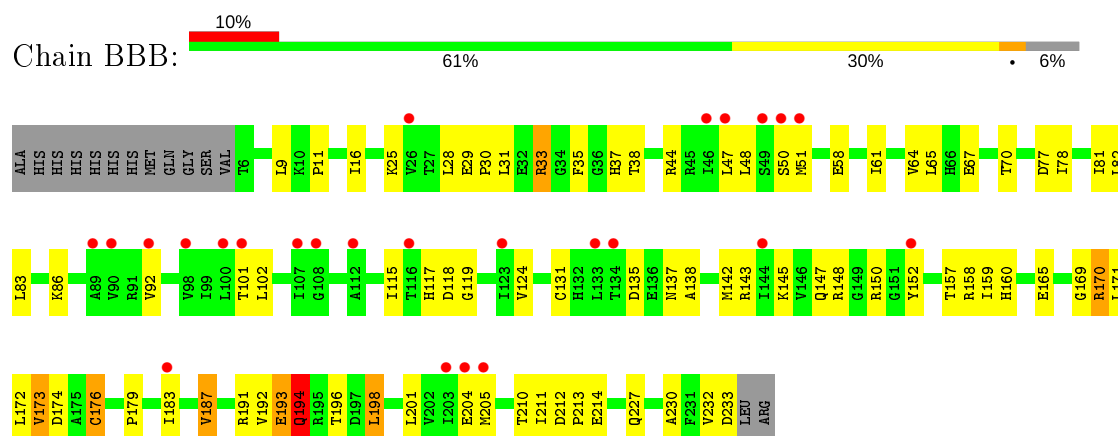
### 3 Residue-property plots [i](#)

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.

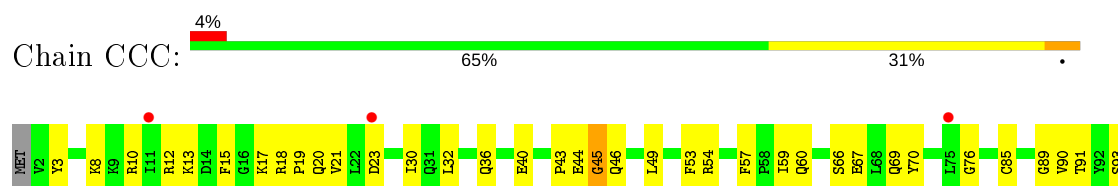
- Molecule 1: DNA-directed RNA polymerase subunit alpha

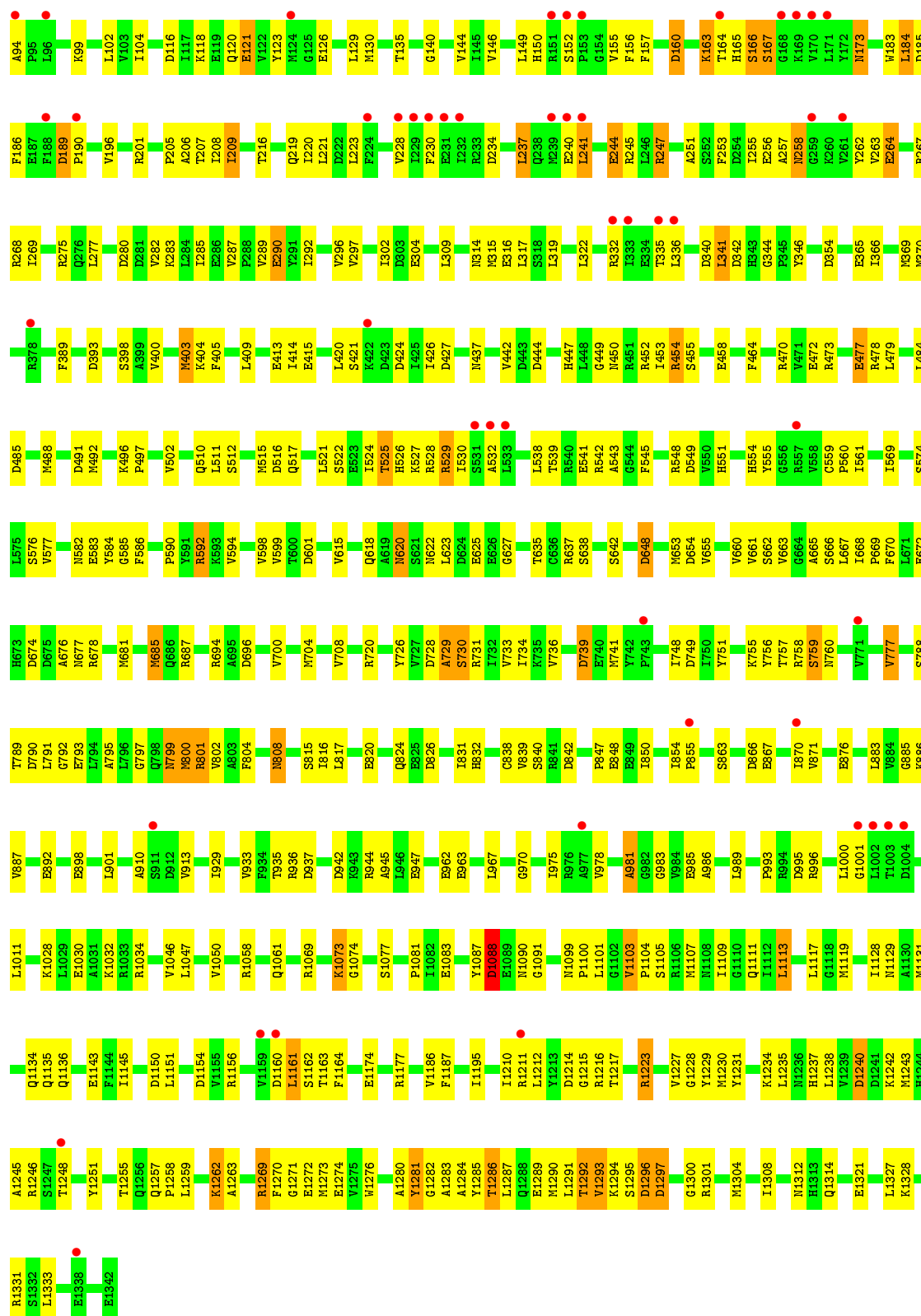


- Molecule 1: DNA-directed RNA polymerase subunit alpha



- Molecule 2: DNA-directed RNA polymerase subunit beta





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





Q1238	G1129	Q1044	V963	S884	E811	G742	G628	L536	E443	R346	L223	A108	MET
R1242	G1130	T1047	K964	V885	D812	M743	F629	T637	G444	V347	L224	A108	LYS
L1243	K1131	T1047	S965	V886	T816	G745	S638	R538	K445	R348	L224	T111	ASP
Q1244	D1133	D1051	V966	S887	E817	M746	V639	S543	Q448	T356	P234	H113	LEU
G1245	L1134	E1052	S969	C888	E818	L746	M644	L544	Q448	T360	M237	H114	LEU
V1246	T1135	L1053	S970	D889	G819	M747	V645	E546	Q448	T360	M237	H115	PHE
N1249	L1138	T1054	L973	T890	I820	A748	M644	V548	D464	L361	L245	K118	LEU
D1250	G1055	G1055	L973	F892	M821	K749	V645	E548	D464	L361	L245	K118	LYS
E1254	E1146	S1058	V974	C893	M821	P750	A659	V549	Q465	R362	P246	P246	ALA
V1255	L1059	S1058	C893	G893	T823	G751	A659	V549	Q465	R362	P246	P246	GLN
I1256	K1151	V1060	T976	V894	T823	G751	A659	V549	Q465	R362	P246	P246	THR
N1257	K1151	V1060	T976	C895	M824	S753	V661	R551	A467	D365	D248	L120	LYS
R1258	L1155	L1062	S977	C898	V825	I754	A662	V552	V468	C366	D248	P121	THR
Q1259	L1156	D1063	L982	C898	I836	P758	E663	V555	L472	G367	D256	S122	THR
M1260	L1156	D1063	K983	K911	G829	I759	G665	V555	L472	G367	D256	S122	E15
L1261	G1161	S1064	L984	K911	G829	I759	G665	V555	L472	G367	D256	S122	E16
I1262	A1065	L985	L985	K911	G829	I759	G665	V555	L472	G367	D256	S122	E16
V1163	V1162	A1065	L985	K911	G829	I759	G665	V555	L472	G367	D256	S122	E16
S1164	V1163	A1065	L985	K911	G829	I759	G665	V555	L472	G367	D256	S122	E16
R1263	F1165	D1073	G989	K992	K992	K992	K992	K992	K992	K992	K992	K992	K992
K1263	F1165	D1073	G989	K992	K992	K992	K992	K992	K992	K992	K992	K992	K992
I1266	E1168	T1169	K1170	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173
G1270	T1169	K1170	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173
S1271	K1170	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173
S1272	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173
L1275	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173
E1276	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173	R1173
Y1282	P1179	D1087	P1091	L1003	V1011	A1097	Q1098	Y1099	E1188	G1103	K1104	A1105	T1106
S1283	D1180	V1186	E1187	E1188	G1103	K1104	A1105	T1106	E1188	G1103	K1104	A1105	T1106
L1284	V1186	E1187	E1188	G1103	K1104	A1105	T1106	E1188	G1103	K1104	A1105	T1106	E1188
V1285	D1181	G1182	S1183	L1196	G1197	V1198	F1199	E1200	G1207	D1208	V1209	S1210	L1211
K1286	G1182	S1183	L1196	G1197	V1198	F1199	E1200	G1207	D1208	V1209	S1210	L1211	L1212
I1287	S1183	L1196	G1197	V1198	F1199	E1200	G1207	D1208	V1209	S1210	L1211	L1212	L1213
R1290	Y1186	E1187	E1188	G1103	K1104	A1105	T1106	E1188	G1103	K1104	A1105	T1106	E1188
E1291	D1181	G1182	S1183	L1196	G1197	V1198	F1199	E1200	G1207	D1208	V1209	S1210	L1211
L1292	V1186	E1187	E1188	G1103	K1104	A1105	T1106	E1188	G1103	K1104	A1105	T1106	E1188
G1296	D1181	G1182	S1183	L1196	G1197	V1198	F1199	E1200	G1207	D1208	V1209	S1210	L1211
K1297	S1183	L1196	G1197	V1198	F1199	E1200	G1207	D1208	V1209	S1210	L1211	L1212	L1213
V1298	E1186	E1187	E1188	G1103	K1104	A1105	T1106	E1188	G1103	K1104	A1105	T1106	E1188
S1303	D1181	G1182	S1183	L1196	G1197	V1198	F1199	E1200	G1207	D1208	V1209	S1210	L1211
L1307	G1207	D1208	V1209	S1210	L1211	L1212	L1213	L1214	L1215	L1216	L1217	L1218	L1219
G1308	D1208	V1209	S1210	L1211	L1212	L1213	L1214	L1215	L1216	L1217	L1218	L1219	L1220
I1309	V1209	S1210	L1211	L1212	L1213	L1214	L1215	L1216	L1217	L1218	L1219	L1220	L1221
T1310	S1210	L1211	L1212	L1213	L1214	L1215	L1216	L1217	L1218	L1219	L1220	L1221	L1222
S1313	L1211	L1212	L1213	L1214	L1215	L1216	L1217	L1218	L1219	L1220	L1221	L1222	L1223
T1316	G1118	D1119	F1037	I1041	D1042	A1122	T1230	F1319	S1324	R101	R220	R345	R535
V1226	G1118	D1119	F1037	I1041	D1042	A1122	T1230	F1319	S1324	R101	R220	R345	R535
H1227	G1118	D1119	F1037	I1041	D1042	A1122	T1230	F1319	S1324	R101	R220	R345	R535
T1230	G1118	D1119	F1037	I1041	D1042	A1122	T1230	F1319	S1324	R101	R220	R345	R535
S1324	G1118	D1119	F1037	I1041	D1042	A1122	T1230	F1319	S1324	R101	R220	R345	R535



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	132.36Å 153.77Å 231.69Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.00 – 5.40 49.02 – 5.40	Depositor EDS
% Data completeness (in resolution range)	98.2 (49.00-5.40) 98.4 (49.02-5.40)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.08 (at 5.39Å)	Xtriage
Refinement program	REFMAC 5.8.0257	Depositor
R, $R_{free}$	0.295 , 0.362 0.282 , 0.351	Depositor DCC
$R_{free}$ test set	770 reflections (4.67%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	225.3	Xtriage
Anisotropy	0.447	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.24 , 293.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.33$ , $\langle L^2 \rangle = 0.17$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.88	EDS
Total number of atoms	29050	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	345.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% 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

### 5.1 Standard geometry

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

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z  >5	RMSZ	# Z  >5
1	AAA	0.64	0/1809	0.73	0/2450
1	BBB	0.65	0/1789	0.73	0/2425
2	CCC	0.63	0/10745	0.78	2/14499 (0.0%)
3	DDD	0.64	0/10729	0.77	0/14487
4	EEE	0.62	0/629	0.76	0/847
5	FFF	0.65	0/2282	0.67	1/3076 (0.0%)
6	111	0.24	0/691	0.63	0/1063
7	222	0.28	0/802	0.64	0/1234
8	333	0.17	0/139	0.59	0/215
All	All	0.62	0/29615	0.75	3/40296 (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

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	-5.93	98.54	110.40
5	FFF	67	TYR	CB-CA-C	5.60	121.59	110.40
2	CCC	454	ARG	CB-CA-C	-5.19	100.01	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

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	1
1	BBB	1767	0	1789	75	0
2	CCC	10576	0	10591	406	0
3	DDD	10568	0	10781	470	5
4	EEE	627	0	634	14	0
5	FFF	2253	0	2298	114	7
6	111	618	0	341	30	3
7	222	716	0	397	29	1
8	333	125	0	66	7	0
9	DDD	2	0	0	2	0
10	333	1	0	0	0	0
10	DDD	1	0	0	0	0
11	DDD	9	0	0	3	0
All	All	29050	0	28710	1042	12

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 18.

The worst 5 of 1042 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:843:VAL:HG11	3:DDD:883:ARG:CB	1.46	1.46
3:DDD:843:VAL:CG1	3:DDD:883:ARG:CB	1.99	1.40
3:DDD:843:VAL:CG1	3:DDD:883:ARG:HB2	1.49	1.40
6:111:54:DA:H2"	6:111:55:DC:C5	1.64	1.30
3:DDD:392:THR:HG21	5:FFF:320:GLN:O	1.27	1.27

The worst 5 of 12 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:CE1	5:FFF:299:ARG:NH2[3_644]	0.94	1.26
5:FFF:67:TYR:CZ	5:FFF:299:ARG:NH2[3_644]	1.07	1.13
5:FFF:67:TYR:CD1	5:FFF:299:ARG:NH2[3_644]	1.75	0.45
5:FFF:67:TYR:CE2	5:FFF:299:ARG:NH2[3_644]	1.89	0.31
5:FFF:67:TYR:CZ	5:FFF:299:ARG:CZ[3_644]	1.98	0.22

## 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%)	214 (94%)	8 (4%)	6 (3%)	5	33
1	BBB	226/242 (93%)	206 (91%)	14 (6%)	6 (3%)	5	33
2	CCC	1339/1342 (100%)	1237 (92%)	77 (6%)	25 (2%)	8	39
3	DDD	1360/1407 (97%)	1253 (92%)	91 (7%)	16 (1%)	13	50
4	EEE	77/90 (86%)	73 (95%)	4 (5%)	0	100	100
5	FFF	275/336 (82%)	256 (93%)	15 (6%)	4 (2%)	10	46
All	All	3505/3659 (96%)	3239 (92%)	209 (6%)	57 (2%)	9	44

5 of 57 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
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%)	181 (91%)	17 (9%)	10	35
1	BBB	196/208 (94%)	182 (93%)	14 (7%)	14	41
2	CCC	1156/1157 (100%)	1069 (92%)	87 (8%)	13	40
3	DDD	1135/1168 (97%)	1064 (94%)	71 (6%)	18	44
4	EEE	67/74 (90%)	64 (96%)	3 (4%)	27	53
5	FFF	240/292 (82%)	230 (96%)	10 (4%)	30	54
All	All	2992/3107 (96%)	2790 (93%)	202 (7%)	16	42

5 of 202 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	CCC	801	ARG
2	CCC	1269	ARG
3	DDD	1309	ILE
2	CCC	815	SER
2	CCC	1073	LYS

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

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	333	5/6 (83%)	2 (40%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
8	333	18	C
8	333	19	U

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 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	10	6,8,8	0.73	0	13,13,13	0.77	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	10	-	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 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	DDD	1504	DPO	3	0

## 5.7 Other polymers

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, 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	AAA	230/242 (95%)	0.24	14 (6%) 21 19	259, 372, 468, 533	0
1	BBB	228/242 (94%)	0.32	25 (10%) 5 8	220, 350, 469, 536	0
2	CCC	1341/1342 (99%)	0.06	52 (3%) 39 33	149, 303, 471, 650	0
3	DDD	1362/1407 (96%)	0.18	97 (7%) 16 15	154, 328, 517, 647	0
4	EEE	79/90 (87%)	-0.17	1 (1%) 77 68	269, 375, 517, 678	0
5	FFF	277/336 (82%)	0.29	13 (4%) 31 28	237, 384, 529, 633	0
6	111	30/50 (60%)	0.62	4 (13%) 3 5	310, 413, 536, 678	0
7	222	35/50 (70%)	0.87	6 (17%) 1 3	271, 388, 687, 743	0
8	333	6/6 (100%)	1.58	0 100 100	331, 366, 398, 405	0
All	All	3588/3765 (95%)	0.16	212 (5%) 22 20	149, 332, 504, 743	0

The worst 5 of 212 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	DDD	941	ALA	9.0
6	111	49	DG	8.3
2	CCC	230	PHE	7.9
3	DDD	1065	ALA	7.2
3	DDD	1098	GLN	6.1

### 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, 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( $\text{\AA}^2$ )	Q<0.9
10	MG	333	101	1/1	0.65	0.31	316,316,316,316	0
11	DPO	DDD	1504	9/9	0.93	0.69	265,283,350,352	0
9	ZN	DDD	1501	1/1	0.96	0.07	509,509,509,509	0
9	ZN	DDD	1502	1/1	0.97	0.14	376,376,376,376	0
10	MG	DDD	1503	1/1	0.97	0.63	477,477,477,477	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.