



wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 10:08 pm BST

PDB ID : 4JK1
Title : X-ray crystal structure of Escherichia coli sigma70 holoenzyme in complex with Guanosine tetraphosphate (ppGpp)
Authors : Murakami, K.S.
Deposited on : 2013-03-09
Resolution : 3.90 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	:	2.11
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)	:	2.11

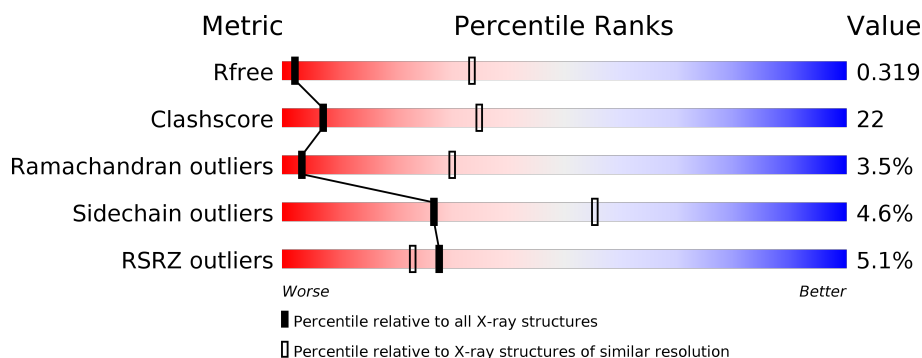
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.90 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	1002 (4.14-3.66)
Clashscore	141614	1004 (4.12-3.68)
Ramachandran outliers	138981	1021 (4.14-3.66)
Sidechain outliers	138945	1014 (4.14-3.66)
RSRZ outliers	127900	1275 (4.20-3.60)

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	A	329	<div> <div>4%</div> <div> <div></div> <div>64%</div> <div>31%</div> <div>••</div> </div> </div>
1	B	329	<div> <div>4%</div> <div> <div></div> <div>39%</div> <div>26%</div> <div>•</div> <div>33%</div> </div> </div>
1	F	329	<div> <div>5%</div> <div> <div></div> <div>48%</div> <div>20%</div> <div>•</div> <div>30%</div> </div> </div>
1	G	329	<div> <div>4%</div> <div> <div></div> <div>39%</div> <div>26%</div> <div>•</div> <div>34%</div> </div> </div>
2	C	1342	<div> <div>3%</div> <div> <div></div> <div>55%</div> <div>39%</div> <div>5%</div> <div>•</div> </div> </div>
2	H	1342	<div> <div>6%</div> <div> <div></div> <div>57%</div> <div>38%</div> <div>5%</div> <div>•</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	D	1407	<div><div><div></div><div></div><div></div><div></div></div><div>3%43%35%5%18%</div></div>
3	I	1407	<div><div><div></div><div></div><div></div><div></div></div><div>5%44%34%•18%</div></div>
4	E	91	<div><div><div></div><div></div><div></div><div></div></div><div>%56%38%••</div></div>
4	J	91	<div><div><div></div><div></div><div></div><div></div></div><div>5%45%36%•16%</div></div>
5	X	613	<div><div><div></div><div></div><div></div><div></div></div><div>7%51%30%•16%</div></div>
5	Y	613	<div><div><div></div><div></div><div></div><div></div></div><div>5%44%28%•25%</div></div>

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 56126 atoms, of which 11 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 Escherichia coli RNA polymerase alpha subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	323	Total	C	N	O	S	0	0	0
			2514	1571	443	492	8			
1	B	221	Total	C	N	O	S	0	0	0
			1706	1065	300	335	6			
1	F	229	Total	C	N	O	S	0	0	0
			1775	1106	313	350	6			
1	G	217	Total	C	N	O	S	0	0	0
			1671	1045	293	327	6			

- Molecule 2 is a protein called Escherichia coli RNA polymerase beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1335	Total	C	N	O	S	0	0	0
			10523	6601	1836	2043	43			
2	H	1335	Total	C	N	O	S	0	0	0
			10523	6601	1836	2043	43			

- Molecule 3 is a protein called Escherichia coli RNA polymerase beta' subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1160	Total	C	N	O	S	0	0	0
			9060	5695	1621	1697	47			
3	I	1160	Total	C	N	O	S	0	0	0
			9060	5695	1621	1697	47			

- Molecule 4 is a protein called Escherichia coli RNA polymerase omega subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			
4	J	76	Total	C	N	O	S	0	0	0
			605	368	115	121	1			

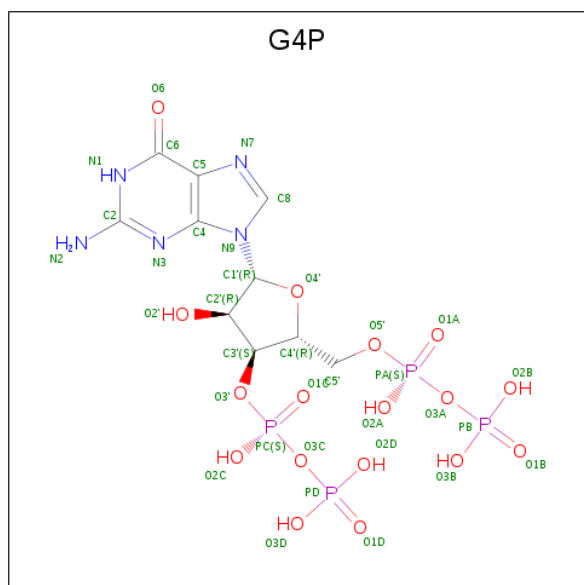
- Molecule 5 is a protein called Escherichia coli RNA polymerase sigma70 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	X	517	Total	C	N	O	S	0	0	0
			4198	2621	745	806	26			
5	Y	458	Total	C	N	O	S	0	0	0
			3732	2335	671	703	23			

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	I	2	Total	Zn	0	0
			2	2		
6	D	2	Total	Zn	0	0
			2	2		

- Molecule 7 is GUANOSINE-5',3'-TETRAPHOSPHATE (three-letter code: G4P) (formula: C₁₀H₁₇N₅O₁₇P₄).

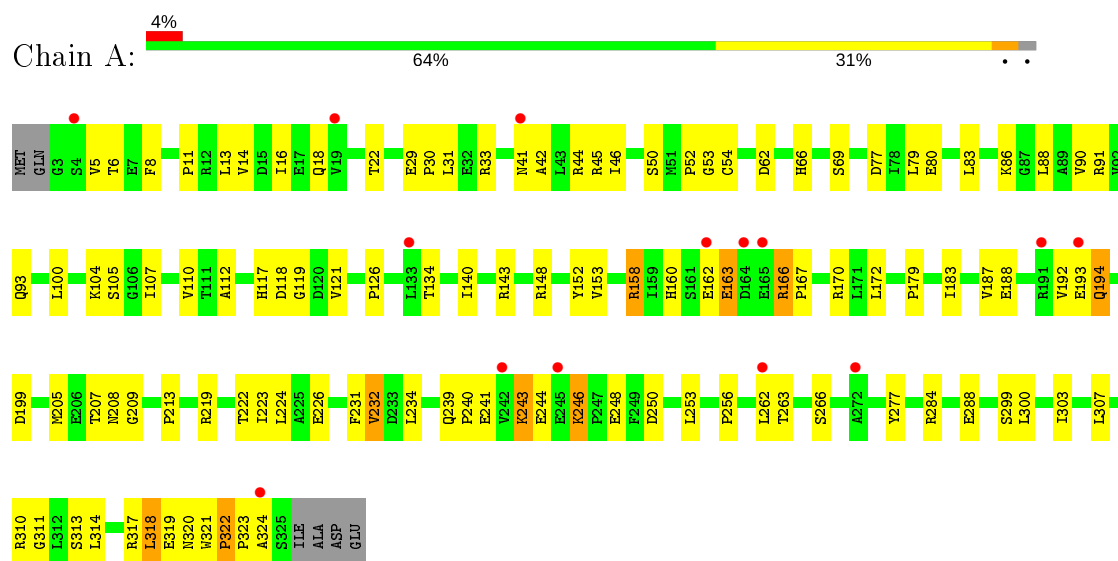


Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	D	1	Total	C	H	N	O	0	0
			47	10	11	5	17		

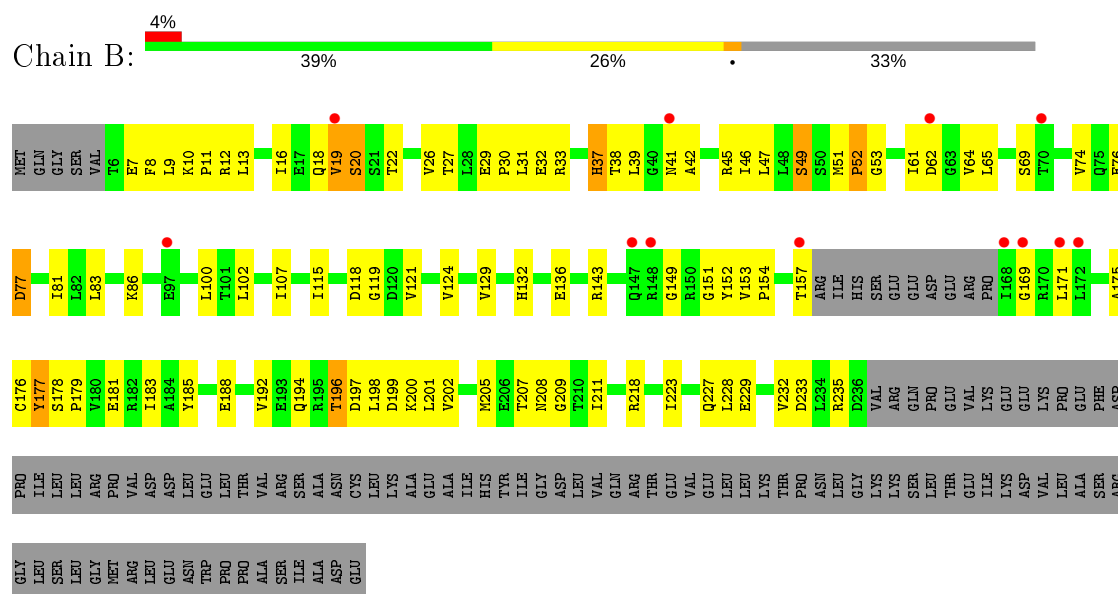
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

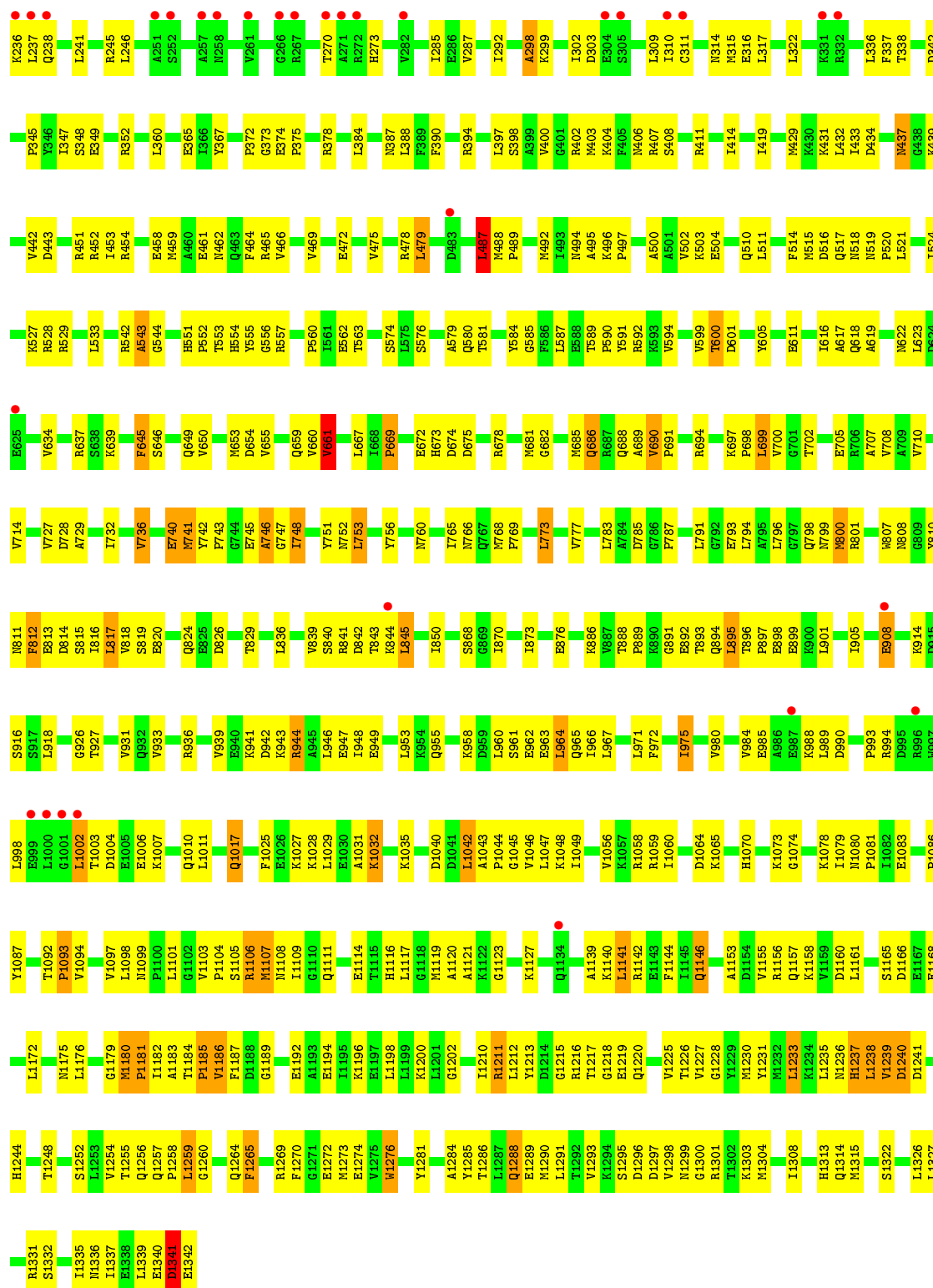
- Molecule 1: Escherichia coli RNA polymerase alpha subunit



- Molecule 1: Escherichia coli RNA polymerase alpha subunit



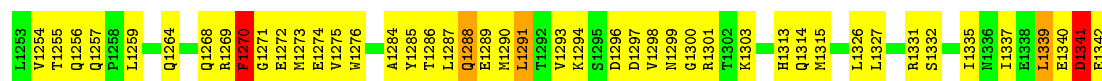
- Molecule 1: Escherichia coli RNA polymerase alpha subunit



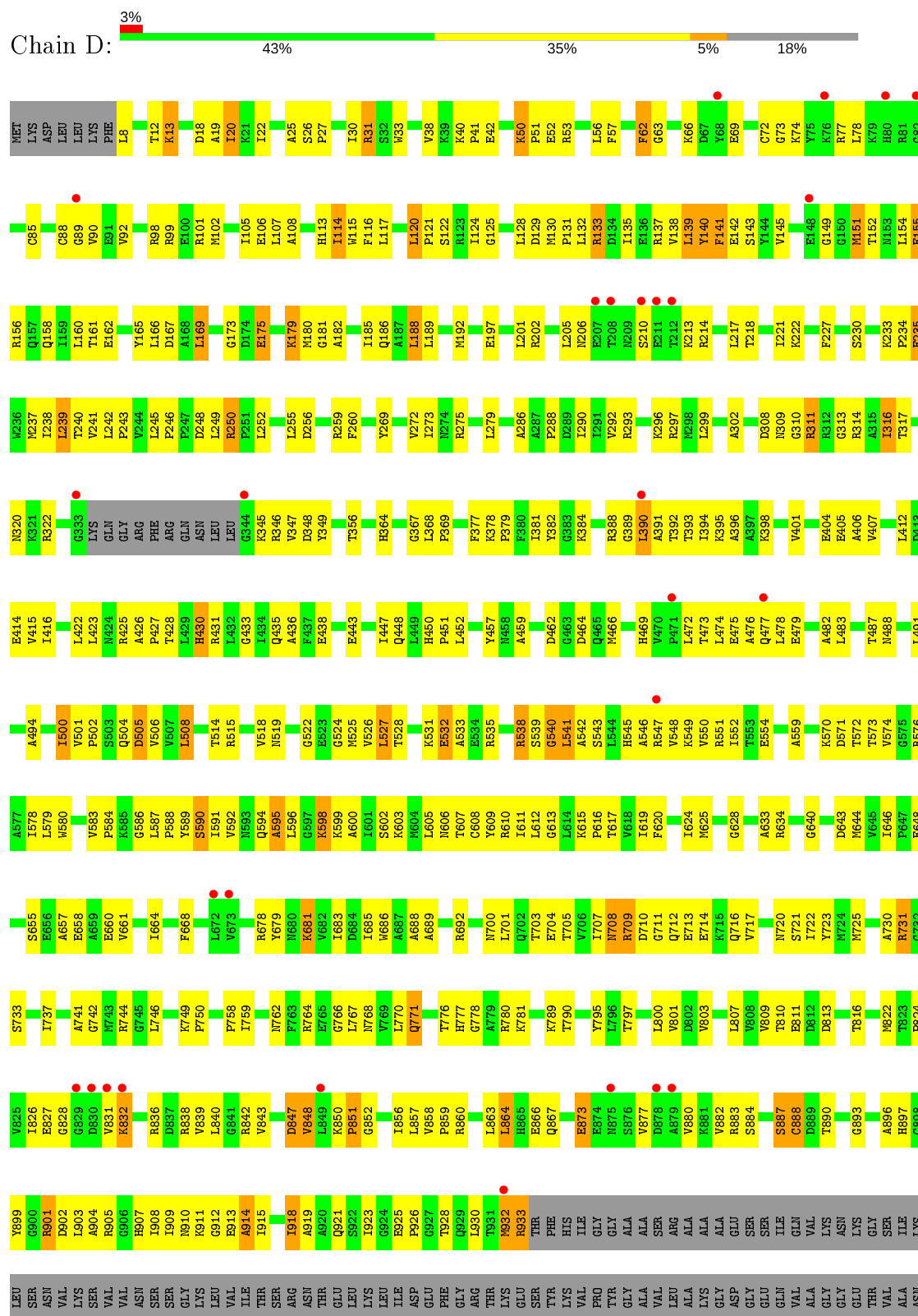
● Molecule 2: Escherichia coli RNA polymerase beta subunit



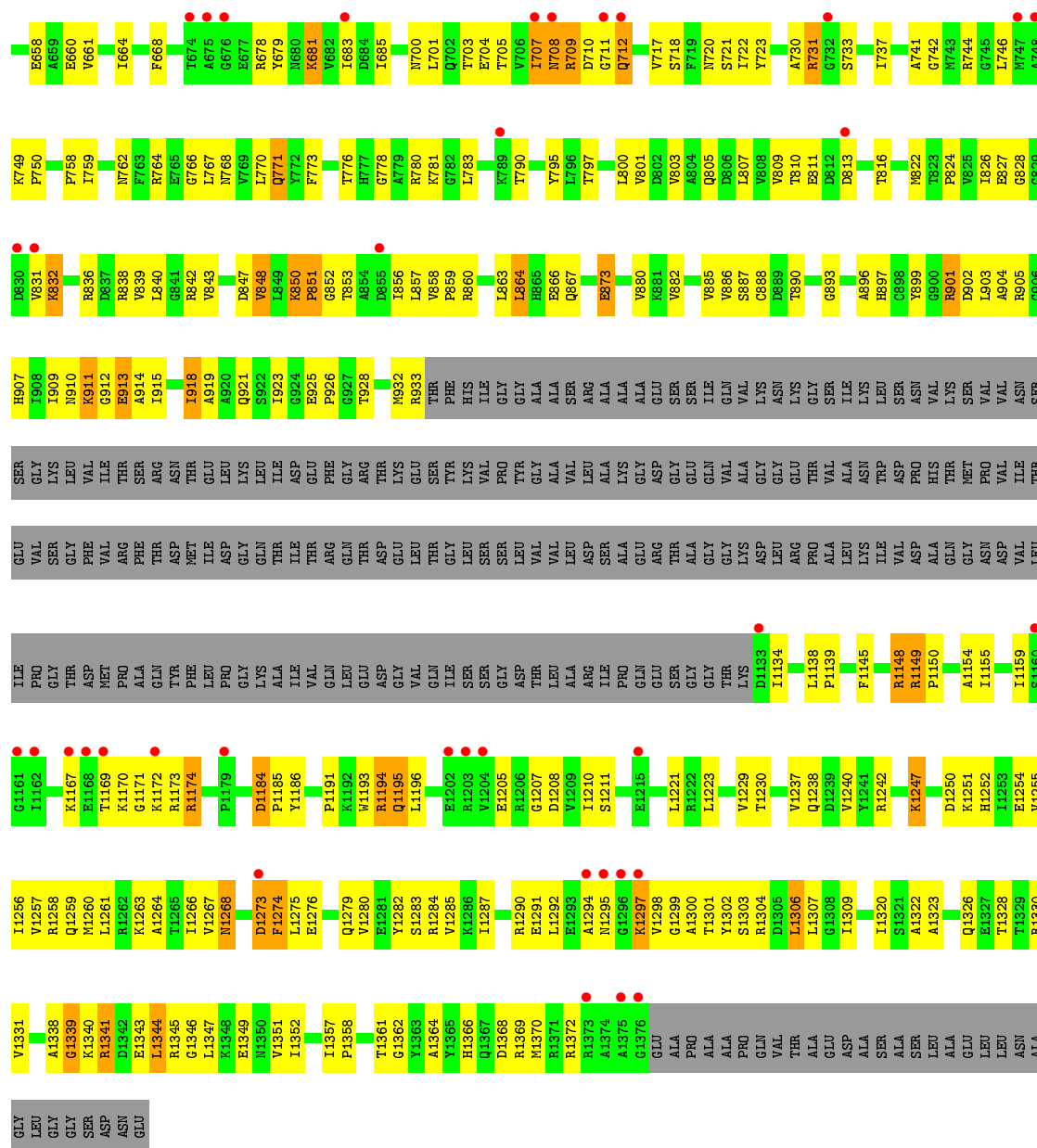
V1169	T1092	D1004	Q932	F812	K639	P552	Q483	M369	A251	R151	S72	MET
P1093	P1094	E1005	V933	E813	F645	T583	F464	G373	S252	S152	Y73	VAL
V1094	D1095	E1006	R936	S815	S646	H554	V469	E374	F253	P153	R74	TYR
I1096	I1097	Q1008	D937	I816	D728	G556	B557	P375	I255	F156	P78	SER
V1097	H1009	H1009	D847	L817	D648	B557	B472	P376	I255	P376	V79	THR
L1098	Q1010	Q1010	Q941	V818	Q649	B557	B473	T377	H258	K161	F80	GLU
M1099	D942	K941	D942	S819	V650	P560	K476	R378	H258	G162		K8
P1100	K943	K943	K943	E820	V650	P560	K476	R378	H258	G162		K9
L1101	R944	R944	R944	T829	M653	T561	B477	L384	Y262	K163	R88	R9
G1102	A945	A945	A945	V735	M653	T561	B477	L384	Y262	K163	R88	R10
L1183	L946	L946	L946	V736	Q658	T563	B478	L384	Y262	K163	R88	R11
T1184	E947	E947	E947	V736	Q659	T563	B478	L384	Y262	K163	R88	R12
S1105	I943	I943	I943	V736	Q659	T563	B478	L384	Y262	K163	R88	R13
P1186	E949	E949	E949	V736	Q659	T563	B478	L384	Y262	K163	R88	R14
V1187				V736	Q659	T563	B478	L384	Y262	K163	R88	R15
F1187	F1025	F1025	F1025	V736	Q659	T563	B478	L384	Y262	K163	R88	R16
D1188	E1026	E1026	E1026	V736	Q659	T563	B478	L384	Y262	K163	R88	R17
G1189	K1027	K1027	K1027	V736	Q659	T563	B478	L384	Y262	K163	R88	R18
E1192	K1028	K1028	K1028	V736	Q659	T563	B478	L384	Y262	K163	R88	R19
A1193	L1029	L1029	L1029	V736	Q659	T563	B478	L384	Y262	K163	R88	R20
E1194	K1032	K1032	K1032	V736	Q659	T563	B478	L384	Y262	K163	R88	R21
I1195	R1033	R1033	R1033	V736	Q659	T563	B478	L384	Y262	K163	R88	R22
T1196	R1034	R1034	R1034	V736	Q659	T563	B478	L384	Y262	K163	R88	R23
E1197	L1042	L1042	L1042	V736	Q659	T563	B478	L384	Y262	K163	R88	R24
L1198	A1043	A1043	A1043	V736	Q659	T563	B478	L384	Y262	K163	R88	R25
L1199	P1044	P1044	P1044	V736	Q659	T563	B478	L384	Y262	K163	R88	R26
K1200	G1045	G1045	G1045	V736	Q659	T563	B478	L384	Y262	K163	R88	R27
L1201	V1046	V1046	V1046	V736	Q659	T563	B478	L384	Y262	K163	R88	R28
G1202	L1047	L1047	L1047	V736	Q659	T563	B478	L384	Y262	K163	R88	R29
Q1209	V1056	V1056	V1056	V736	Q659	T563	B478	L384	Y262	K163	R88	R30
I1210	K1057	K1057	K1057	V736	Q659	T563	B478	L384	Y262	K163	R88	R31
R1211	R1058	R1058	R1058	V736	Q659	T563	B478	L384	Y262	K163	R88	R32
G1215	R1059	R1059	R1059	V736	Q659	T563	B478	L384	Y262	K163	R88	R33
R1216	D1064	D1064	D1064	V736	Q659	T563	B478	L384	Y262	K163	R88	R34
T1217	K1065	K1065	K1065	V736	Q659	T563	B478	L384	Y262	K163	R88	R35
G1218	R1069	R1069	R1069	V736	Q659	T563	B478	L384	Y262	K163	R88	R36
E1219	K1070	K1070	K1070	V736	Q659	T563	B478	L384	Y262	K163	R88	R37
R1223	H1071	H1071	H1071	V736	Q659	T563	B478	L384	Y262	K163	R88	R38
P1224	G1072	G1072	G1072	V736	Q659	T563	B478	L384	Y262	K163	R88	R39
V1225	K1073	K1073	K1073	V736	Q659	T563	B478	L384	Y262	K163	R88	R40
T1226	G1074	G1074	G1074	V736	Q659	T563	B478	L384	Y262	K163	R88	R41
V1227	K1078	K1078	K1078	V736	Q659	T563	B478	L384	Y262	K163	R88	R42
G1228	I1079	I1079	I1079	V736	Q659	T563	B478	L384	Y262	K163	R88	R43
Y1231	M1080	M1080	M1080	V736	Q659	T563	B478	L384	Y262	K163	R88	R44
M1232	P1081	P1081	P1081	V736	Q659	T563	B478	L384	Y262	K163	R88	R45
L1233	I1082	I1082	I1082	V736	Q659	T563	B478	L384	Y262	K163	R88	R46
M1236	D995	D995	D995	V736	Q659	T563	B478	L384	Y262	K163	R88	R47
E1237	R996	R996	R996	V736	Q659	T563	B478	L384	Y262	K163	R88	R48
K1238	V913	V913	V913	V736	Q659	T563	B478	L384	Y262	K163	R88	R49
V1239	H997	H997	H997	V736	Q659	T563	B478	L384	Y262	K163	R88	R50
L1240	L998	L998	L998	V736	Q659	T563	B478	L384	Y262	K163	R88	R51
D1241	P1086	P1086	P1086	V736	Q659	T563	B478	L384	Y262	K163	R88	R52
S1252	V1087	V1087	V1087	V736	Q659	T563	B478	L384	Y262	K163	R88	R53
	D1088	D1088	D1088	V736	Q659	T563	B478	L384	Y262	K163	R88	R54
	E1089	E1089	E1089	V736	Q659	T563	B478	L384	Y262	K163	R88	R55
	S1165	S1165	S1165	V736	Q659	T563	B478	L384	Y262	K163	R88	R56
	D1166	D1166	D1166	V736	Q659	T563	B478	L384	Y262	K163	R88	R57
				V736	Q659	T563	B478	L384	Y262	K163	R88	R58
				V736	Q659	T563	B478	L384	Y262	K163	R88	R59
				V736	Q659	T563	B478	L384	Y262	K163	R88	R60
				V736	Q659	T563	B478	L384	Y262	K163	R88	R61
				V736	Q659	T563	B478	L384	Y262	K163	R88	R62
				V736	Q659	T563	B478	L384	Y262	K163	R88	R63
				V736	Q659	T563	B478	L384	Y262	K163	R88	R64
				V736	Q659	T563	B478	L384	Y262	K163	R88	R65
				V736	Q659	T563	B478	L384	Y262	K163	R88	R66
				V736	Q659	T563	B478	L384	Y262	K163	R88	R67
				V736	Q659	T563	B478	L384	Y262	K163	R88	R68
				V736	Q659	T563	B478	L384	Y262	K163	R88	R69
				V736	Q659	T563	B478	L384	Y262	K163	R88	R70
				V736	Q659	T563	B478	L384	Y262	K163	R88	R71
				V736	Q659	T563	B478	L384	Y262	K163	R88	R72
				V736	Q659	T563	B478	L384	Y262	K163	R88	R73
				V736	Q659	T563	B478	L384	Y262	K163	R88	R74
				V736	Q659	T563	B478	L384	Y262	K163	R88	R75
				V736	Q659	T563	B478	L384	Y262	K163	R88	R76
				V736	Q659	T563	B478	L384	Y262	K163	R88	R77
				V736	Q659	T563	B478	L384	Y262	K163	R88	R78
				V736	Q659	T563	B478	L384	Y262	K163	R88	R79
				V736	Q659	T563	B478	L384	Y262	K163	R88	R80
				V736	Q659	T563	B478	L384	Y262	K163	R88	R81
				V736	Q659	T563	B478	L384	Y262	K163	R88	R82
				V736	Q659	T563	B478	L384	Y262	K163	R88	R83
				V736	Q659	T563	B478	L384	Y262	K163	R88	R84
				V736	Q659	T563	B478	L384	Y262	K163	R88	R85
				V736	Q659	T563	B478	L384	Y262	K163	R88	R86
				V736	Q659	T563	B478	L384	Y262	K163	R88	R87
				V736	Q659	T563	B478	L384	Y262	K163	R88	R88
				V736	Q659	T563	B478	L384	Y262	K163	R88	R89
				V736	Q659	T563	B478	L384	Y262	K163	R88	R90
				V736	Q659	T563	B478	L384	Y262	K163	R88	R91
				V736	Q659	T563	B478	L384	Y262	K163	R88	R92
				V736	Q659	T563	B478	L384	Y262	K163	R88	R93
				V736	Q659	T563	B478	L384	Y262	K163	R88	R94
				V736	Q659	T563	B478	L384	Y262	K163	R88	R95
				V736	Q659	T563	B478	L384	Y262	K163	R88	R96
				V736	Q659	T563	B478	L384	Y262	K163	R88	R97
				V736	Q659	T563	B478	L384	Y262	K163	R88	R98
				V736	Q659	T563	B478	L384	Y262	K163	R88	R99
				V736	Q659	T563	B478	L384	Y262	K163	R88	R100
				V736	Q659	T563	B478	L384	Y262	K163	R88	R101
				V736	Q659	T563	B478	L384	Y262	K163	R88	R102
				V736	Q659	T563	B478	L384	Y262	K163	R88	R103
				V736	Q659	T563	B478	L384	Y262	K163	R88	R104
				V736	Q659	T563	B478	L384	Y262	K163	R88	R105
				V736	Q659	T563	B478	L384	Y262	K163	R88	R106
				V736	Q659	T563	B478	L384	Y262	K163	R88	R107
				V736	Q659	T563	B478	L384	Y262	K163	R88	R108
				V736	Q659	T563	B478	L384	Y262	K163	R88	R109
				V736	Q659	T563	B478	L384	Y262	K163	R88	R110
				V736	Q659	T563	B478	L384	Y262	K163	R88	R111
				V736	Q659	T563	B478	L384	Y262	K163	R88	R112
				V736	Q659	T563	B478	L384	Y262	K163	R88	R113
				V736	Q659	T563	B478	L384	Y262	K163	R88	R114
				V736	Q659	T563	B478	L384	Y262	K163	R88	R115
				V736	Q659	T563	B478	L384	Y262	K163	R88	R116
				V736	Q659	T563	B478	L384	Y262	K163	R88	R117
				V736	Q659	T563	B478	L384	Y262	K163	R88	R118
				V736	Q659	T563	B478	L384	Y262	K163	R88	R119



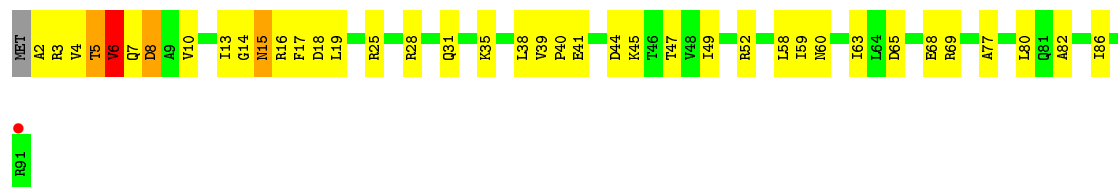
● Molecule 3: Escherichia coli RNA polymerase beta' subunit







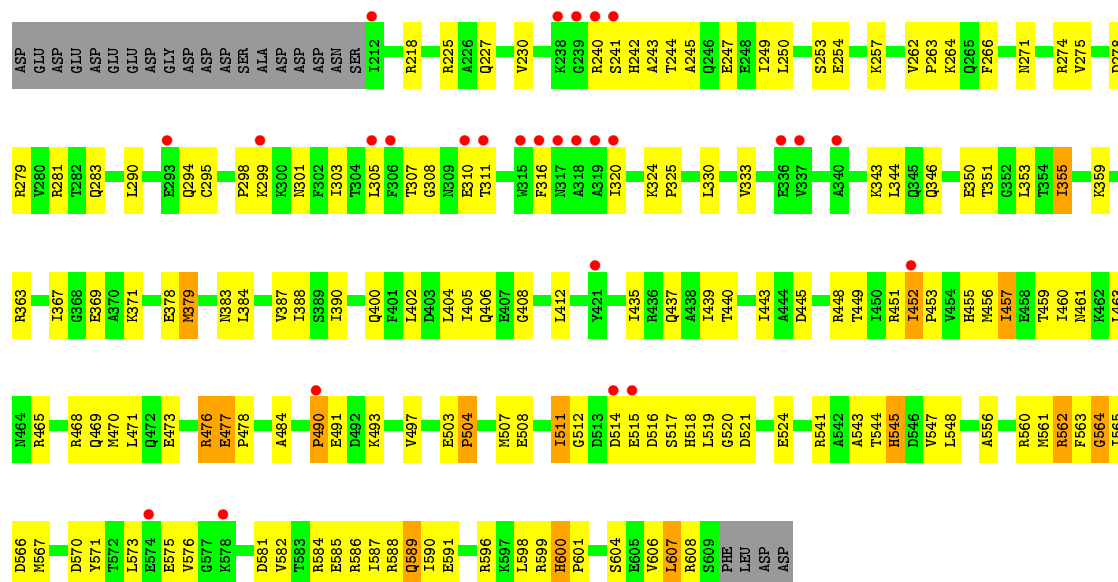
- Molecule 4: Escherichia coli RNA polymerase omega subunit



- Molecule 4: Escherichia coli RNA polymerase omega subunit







4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	184.57Å 203.82Å 307.67Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.84 – 3.90 29.84 – 3.85	Depositor EDS
% Data completeness (in resolution range)	89.6 (29.84-3.90) 82.4 (29.84-3.85)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.78 (at 3.86Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, R_{free}	0.252 , 0.320 0.253 , 0.319	Depositor DCC
R_{free} test set	4799 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	124.7	Xtriage
Anisotropy	0.149	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 37.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	56126	wwPDB-VP
Average B, all atoms (Å ²)	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.73% 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: G4P, 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	A	0.19	0/2548	0.37	0/3454
1	B	0.19	0/1725	0.38	0/2337
1	F	0.19	0/1797	0.38	0/2436
1	G	0.19	0/1690	0.37	0/2290
2	C	0.20	0/10690	0.37	0/14423
2	H	0.20	0/10690	0.37	0/14423
3	D	0.20	0/9198	0.38	0/12413
3	I	0.20	0/9198	0.38	0/12413
4	E	0.19	0/710	0.38	0/956
4	J	0.19	0/607	0.37	0/817
5	X	0.20	0/4253	0.36	0/5719
5	Y	0.20	0/3783	0.35	0/5083
All	All	0.20	0/56889	0.37	0/76764

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	A	2514	0	2566	97	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1706	0	1738	89	0
1	F	1775	0	1800	62	0
1	G	1671	0	1706	84	0
2	C	10523	0	10546	517	0
2	H	10523	0	10546	501	0
3	D	9060	0	9257	530	0
3	I	9060	0	9257	511	0
4	E	708	0	719	40	0
4	J	605	0	612	32	0
5	X	4198	0	4250	169	0
5	Y	3732	0	3809	137	0
6	D	2	0	0	0	0
6	I	2	0	0	0	0
7	D	36	11	11	2	0
All	All	56115	11	56817	2538	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 22.

The worst 5 of 2538 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:I:1173:ARG:HA	3:I:1174:ARG:HB2	1.28	1.14
3:D:1173:ARG:HA	3:D:1174:ARG:HB2	1.29	1.11
3:D:310:GLY:HA3	3:D:311:ARG:HB2	1.23	1.11
2:H:488:MET:HB2	2:H:490:GLN:H	1.14	1.05
2:C:42:ASP:HB3	2:C:43:PRO:HD2	1.38	1.02

There are no symmetry-related clashes.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	321/329 (98%)	266 (83%)	41 (13%)	14 (4%)	2	25
1	B	217/329 (66%)	186 (86%)	24 (11%)	7 (3%)	4	32
1	F	227/329 (69%)	196 (86%)	26 (12%)	5 (2%)	6	38
1	G	213/329 (65%)	188 (88%)	21 (10%)	4 (2%)	8	41
2	C	1333/1342 (99%)	1069 (80%)	213 (16%)	51 (4%)	3	28
2	H	1333/1342 (99%)	1070 (80%)	213 (16%)	50 (4%)	3	28
3	D	1154/1407 (82%)	922 (80%)	189 (16%)	43 (4%)	3	29
3	I	1154/1407 (82%)	929 (80%)	183 (16%)	42 (4%)	3	29
4	E	88/91 (97%)	76 (86%)	8 (9%)	4 (4%)	2	25
4	J	74/91 (81%)	64 (86%)	6 (8%)	4 (5%)	2	22
5	X	511/613 (83%)	450 (88%)	45 (9%)	16 (3%)	4	32
5	Y	454/613 (74%)	411 (90%)	32 (7%)	11 (2%)	6	37
All	All	7079/8222 (86%)	5827 (82%)	1001 (14%)	251 (4%)	3	30

5 of 251 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	319	GLU
1	B	20	SER
1	B	52	PRO
2	C	21	VAL
2	C	39	ILE

5.3.2 Protein sidechains ⓘ

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	A	281/286 (98%)	274 (98%)	7 (2%)	47	69
1	B	189/286 (66%)	185 (98%)	4 (2%)	53	73
1	F	197/286 (69%)	192 (98%)	5 (2%)	47	69
1	G	185/286 (65%)	180 (97%)	5 (3%)	44	67

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	C	1150/1157 (99%)	1087 (94%)	63 (6%)	21	51
2	H	1150/1157 (99%)	1092 (95%)	58 (5%)	24	53
3	D	971/1168 (83%)	919 (95%)	52 (5%)	22	52
3	I	971/1168 (83%)	921 (95%)	50 (5%)	24	53
4	E	74/75 (99%)	71 (96%)	3 (4%)	30	58
4	J	65/75 (87%)	65 (100%)	0	100	100
5	X	460/540 (85%)	442 (96%)	18 (4%)	32	59
5	Y	407/540 (75%)	391 (96%)	16 (4%)	32	59
All	All	6100/7024 (87%)	5819 (95%)	281 (5%)	27	55

5 of 281 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	E	8	ASP
2	H	9	LYS
3	I	1247	LYS
5	X	28	ASN
5	X	476	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 84 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	1268	ASN
5	X	469	GLN
5	Y	242	HIS
4	E	31	GLN
5	X	54	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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 [i](#)

There are no carbohydrates 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	$\# Z > 2$
7	G4P	D	1503	-	30,38,38	1.86	7 (23%)	43,61,61	1.77	9 (20%)

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
7	G4P	D	1503	-	-	7/23/43/43	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	D	1503	G4P	C2'-C1'	-5.25	1.45	1.53
7	D	1503	G4P	C2-N2	4.19	1.42	1.33
7	D	1503	G4P	C2'-C3'	-3.55	1.45	1.52
7	D	1503	G4P	C5-C4	-3.42	1.31	1.40
7	D	1503	G4P	C8-N7	2.37	1.38	1.34

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	D	1503	G4P	O3C-PC-O3'	4.89	112.34	102.48
7	D	1503	G4P	O4'-C1'-C2'	-4.68	100.09	106.93
7	D	1503	G4P	PC-O3C-PD	-4.57	117.16	132.83
7	D	1503	G4P	PA-O3A-PB	-3.42	121.10	132.83

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	D	1503	G4P	C5-C6-N1	-3.26	118.97	123.43

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

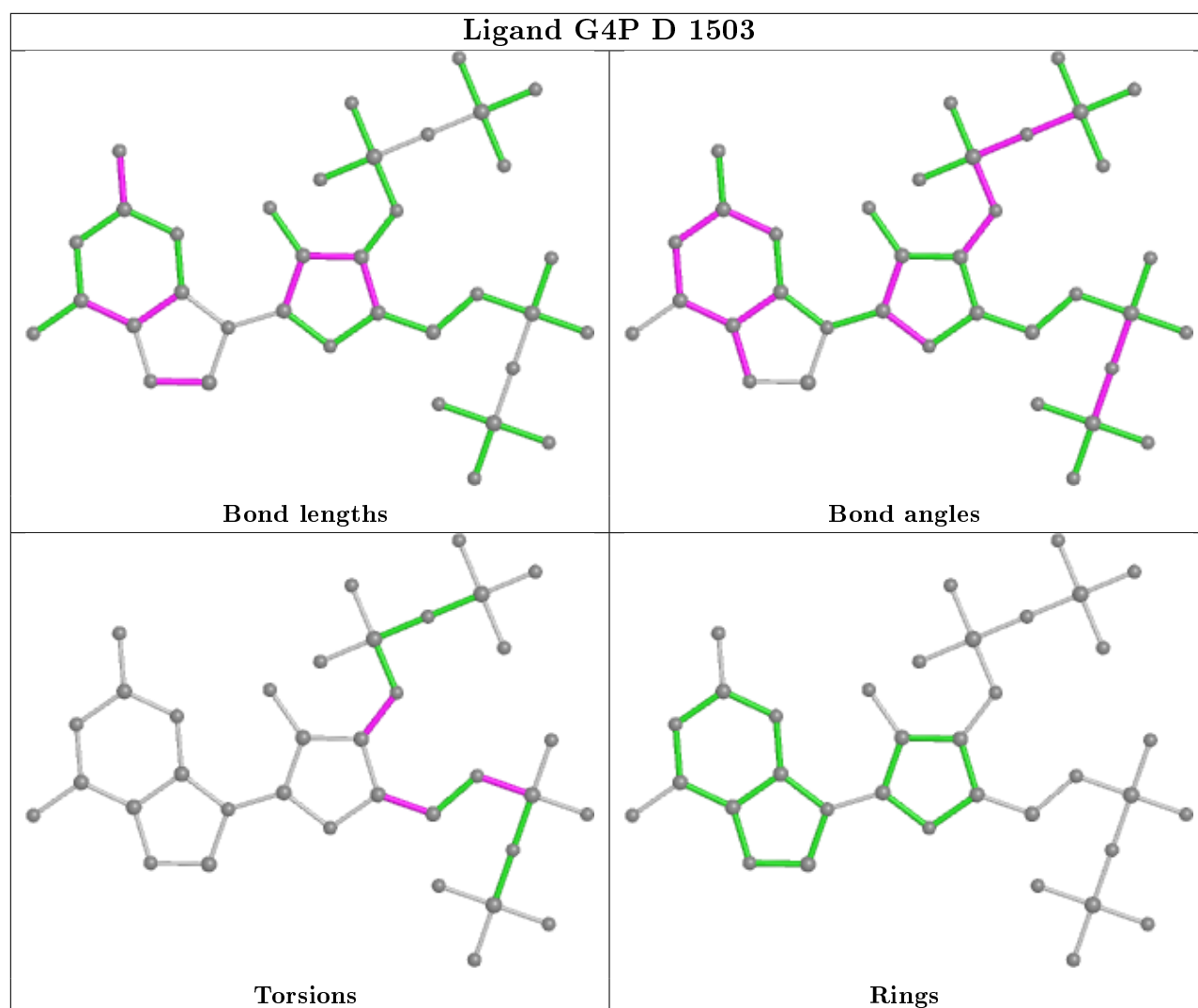
Mol	Chain	Res	Type	Atoms
7	D	1503	G4P	C5'-O5'-PA-O1A
7	D	1503	G4P	C5'-O5'-PA-O2A
7	D	1503	G4P	O4'-C4'-C5'-O5'
7	D	1503	G4P	C3'-C4'-C5'-O5'
7	D	1503	G4P	C4'-C3'-O3'-PC

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	D	1503	G4P	2	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less than 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

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

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	A	323/329 (98%)	-0.01	14 (4%)	35	28	1, 73, 178, 299	0
1	B	221/329 (67%)	0.29	12 (5%)	25	21	5, 98, 204, 259	0
1	F	229/329 (69%)	0.30	15 (6%)	18	13	27, 123, 199, 278	0
1	G	217/329 (65%)	0.34	13 (5%)	21	16	34, 118, 187, 236	0
2	C	1335/1342 (99%)	-0.15	42 (3%)	49	38	0, 48, 170, 262	0
2	H	1335/1342 (99%)	0.09	80 (5%)	21	16	0, 88, 203, 293	0
3	D	1160/1407 (82%)	-0.04	42 (3%)	42	33	0, 42, 158, 289	0
3	I	1160/1407 (82%)	0.09	71 (6%)	21	15	0, 60, 192, 317	0
4	E	90/91 (98%)	-0.27	1 (1%)	80	73	1, 52, 121, 158	0
4	J	76/91 (83%)	0.25	5 (6%)	18	13	10, 89, 165, 211	0
5	X	517/613 (84%)	0.10	42 (8%)	12	9	0, 101, 226, 326	0
5	Y	458/613 (74%)	0.14	29 (6%)	20	14	1, 109, 234, 357	0
All	All	7121/8222 (86%)	0.04	366 (5%)	28	23	0, 73, 197, 357	0

The worst 5 of 366 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	10	ALA	12.6
5	Y	239	GLY	9.4
3	I	212	THR	9.0
3	I	11	GLN	8.3
3	I	1161	GLY	7.6

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 carbohydrates in this entry.

6.4 Ligands [i](#)

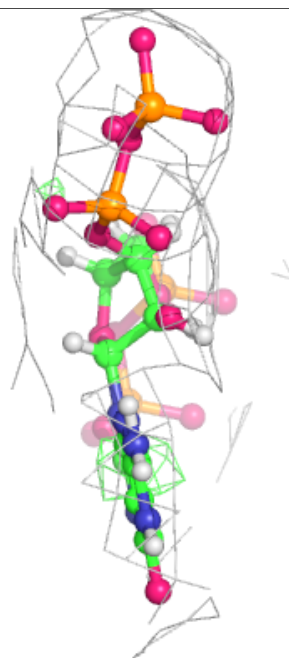
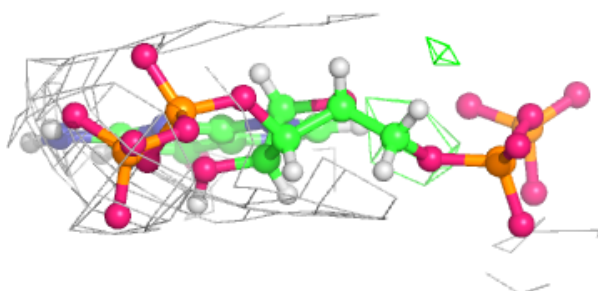
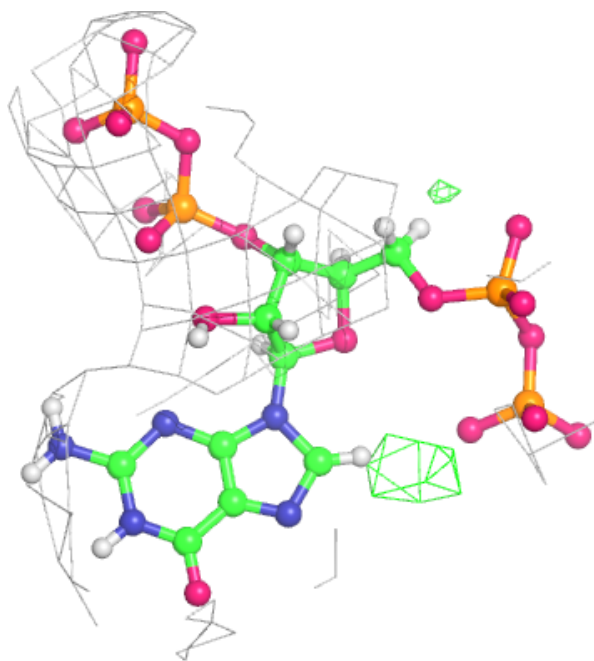
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
7	G4P	D	1503	36/36	0.83	0.20	31,56,93,118	0
6	ZN	I	1501	1/1	0.97	0.07	60,60,60,60	0
6	ZN	I	1502	1/1	0.98	0.15	49,49,49,49	0
6	ZN	D	1501	1/1	0.98	0.08	54,54,54,54	0
6	ZN	D	1502	1/1	0.99	0.18	8,8,8,8	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 G4P D 1503:

$2mF_o-DF_c$ (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.