



# wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 06:40 PM GMT

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 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

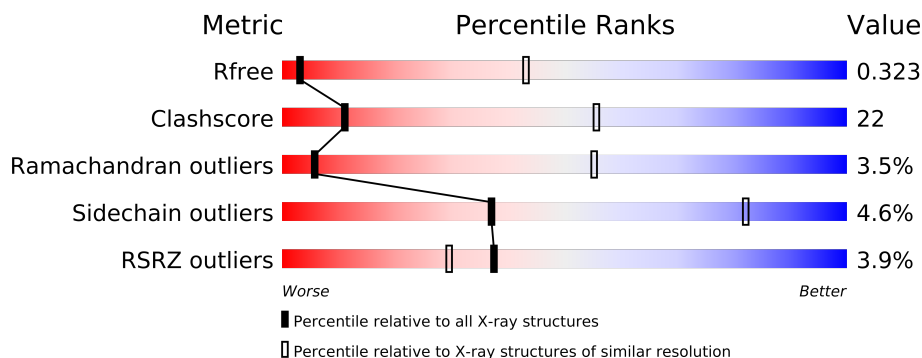
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

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))
$R_{free}$	66092	1022 (4.38-3.42)
Clashscore	79885	1173 (4.30-3.50)
Ramachandran outliers	78287	1118 (4.30-3.50)
Sidechain outliers	78261	1107 (4.30-3.50)
RSRZ outliers	66119	1000 (4.36-3.44)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	329	
1	B	329	
1	F	329	
1	G	329	
2	C	1342	
2	H	1342	
3	D	1407	
3	I	1407	
4	E	91	
4	J	91	
5	X	613	
5	Y	613	

## 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 deuterium.

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			

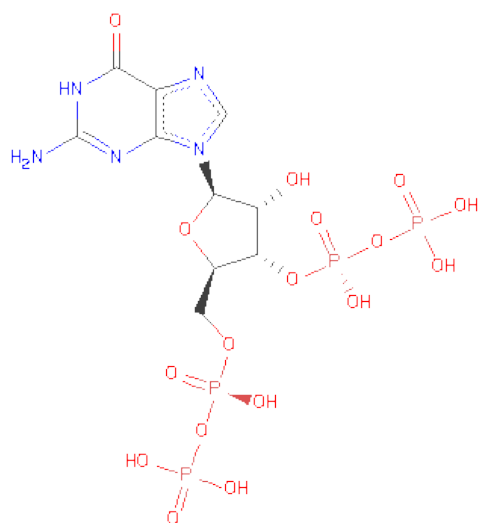
- Molecule 5 is a protein called Escherichia coli RNA polymerase sigma70 subunit.

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).

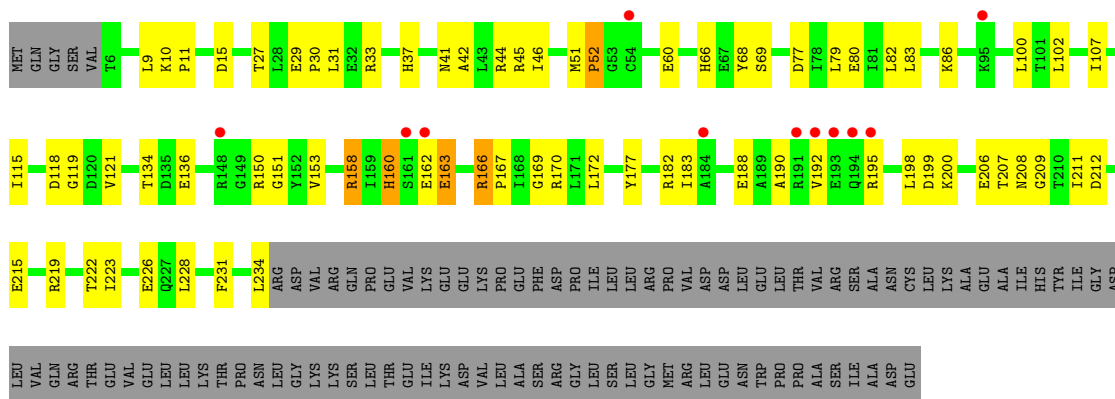
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	I	2	Total	Zn	0	0
			2	2		
6	D	2	Total	Zn	0	0
			2	2		

- Molecule 7 is GUANOSINE-5',3'-TETRAPHOSPHATE (three-letter code: G4P) (formula: C<sub>10</sub>H<sub>17</sub>N<sub>5</sub>O<sub>17</sub>P<sub>4</sub>).



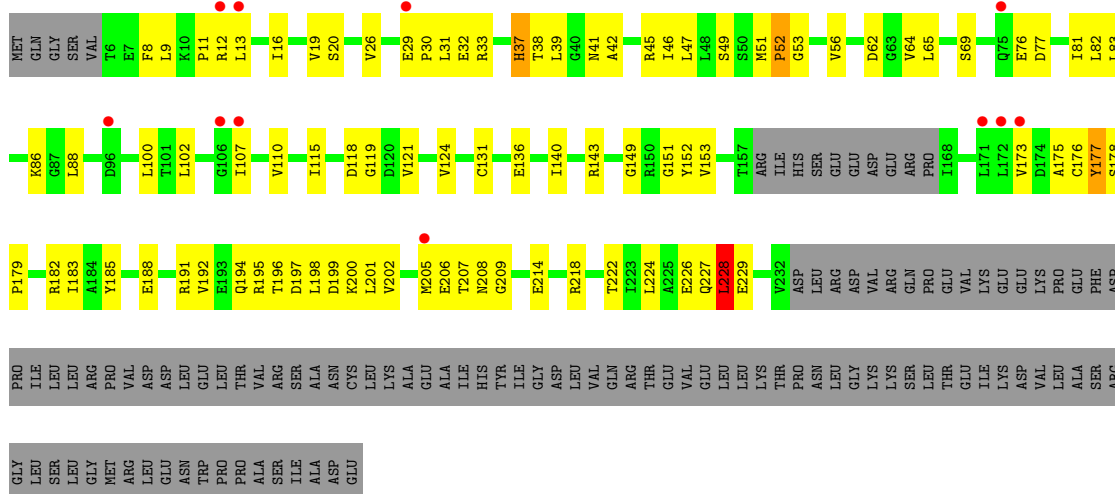
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	D	1	Total	C	H	N	O	0	0
			47	10	11	5	17		





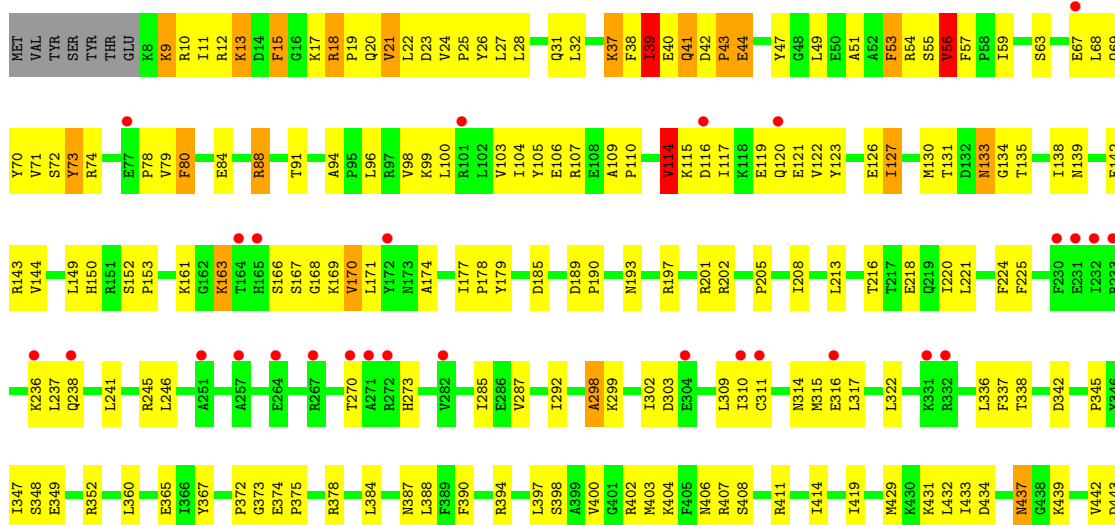
• Molecule 1: Escherichia coli RNA polymerase alpha subunit

Chain G:



• Molecule 2: Escherichia coli RNA polymerase beta subunit

Chain C:



MET	VAL	SF2	S152	MET	VAL	SF2	S152
	TVR	Y73	P153		TVR	Y73	P153
	SER	R74	F156		SER	R74	F156
	TVR	P78	K161		TVR	P78	K161
	THR	V79	G162		THR	V79	G162
	GLU	F80	K163		GLU	F80	K163
	K8	R88	T164		K8	R88	T164
	K9	T91	S166		K9	T91	S166
	R10	Y92	S167		R10	Y92	S167
	I11	S93	G168		I11	S93	G168
R12	A94	K169	R12	A94	K169		
K13	V98	V170	K13	V98	V170		
D14	K99	L171	D14	K99	L171		
F15	L100	R175	F15	L100	R175		
G16	R101	I176	G16	R101	I176		
G17	L102	I177	G17	L102	I177		
K17	V103	P178	K17	V103	P178		
R18	I104	Y179	R18	I104	Y179		
P19	E105	L184	P19	E105	L184		
D20	Y106	D185	D20	Y106	D185		
D23	E107	P189	D23	E107	P189		
D23	E108	P190	D23	E108	P190		
V24	R107	N193	V24	R107	N193		
P25	E109	R197	P25	E109	R197		
P25	P110	R202	P25	P110	R202		
L27	P111	L204	L27	P111	L204		
L28	V114	P205	L28	V114	P205		
Q36	K37	E119	Q36	K37	E119		
	F38	Q120		F38	Q120		
	I39	E121		I39	E121		
	D42	V122		D42	V122		
	P43	Y123		P43	Y123		
	E44	E126		E44	E126		
	G45	I127		G45	I127		
	Q46	P128		Q46	P128		
	Y47	L129		Y47	L129		
	G48	M130		G48	M130		
L49	T131	L49	T131				
F53	D132	F53	D132				
R54	M133	R54	M133				
S55	G134	S55	G134				
Y56	T135	Y56	T135				
F57	I138	F57	I138				
P58	M139	P58	M139				
I59	E142	I59	E142				
Q60	R143	Q60	R143				
S81	Y144	S81	Y144				
Y62	I145	Y62	I145				
N65	Q148	N65	Q148				
L68	L149	L68	L149				
Q69	H150	Q69	H150				
Y70	P151	Y70	P151				
W71	T155	W71	T155				
R256	V262	K236	R256	V262	K236		
	Y263	L237		Y263	L237		
	E264	Q238		E264	Q238		
	R267	R368		R267	R368		
	H273	L241		H273	L241		
	L285	R245		L285	R245		
	L292	L246		L292	L246		
	V297	G373		V297	G373		
	A298	E374		A298	E374		
	K299	P375		K299	P375		
T302	MET	T302	MET				
D303		D303					
L309		L309					
C311		C311					
L317		L317					
L322		L322					
R332		R332					
I333		I333					
E334		E334					
T335		T335					
L336	L336						
F337	F337						
D342	D342						
H343	H343						
G344	G344						
P345	P345						
Y346	Y346						
L347	L347						
T356	T356						
L360	L360						
S361	S361						
A362	A362						
G365	G365						
L366	L366						
Y367	Y367						
R368	R368						
M369	M369						
G373	G373						
E374	E374						
P375	P375						

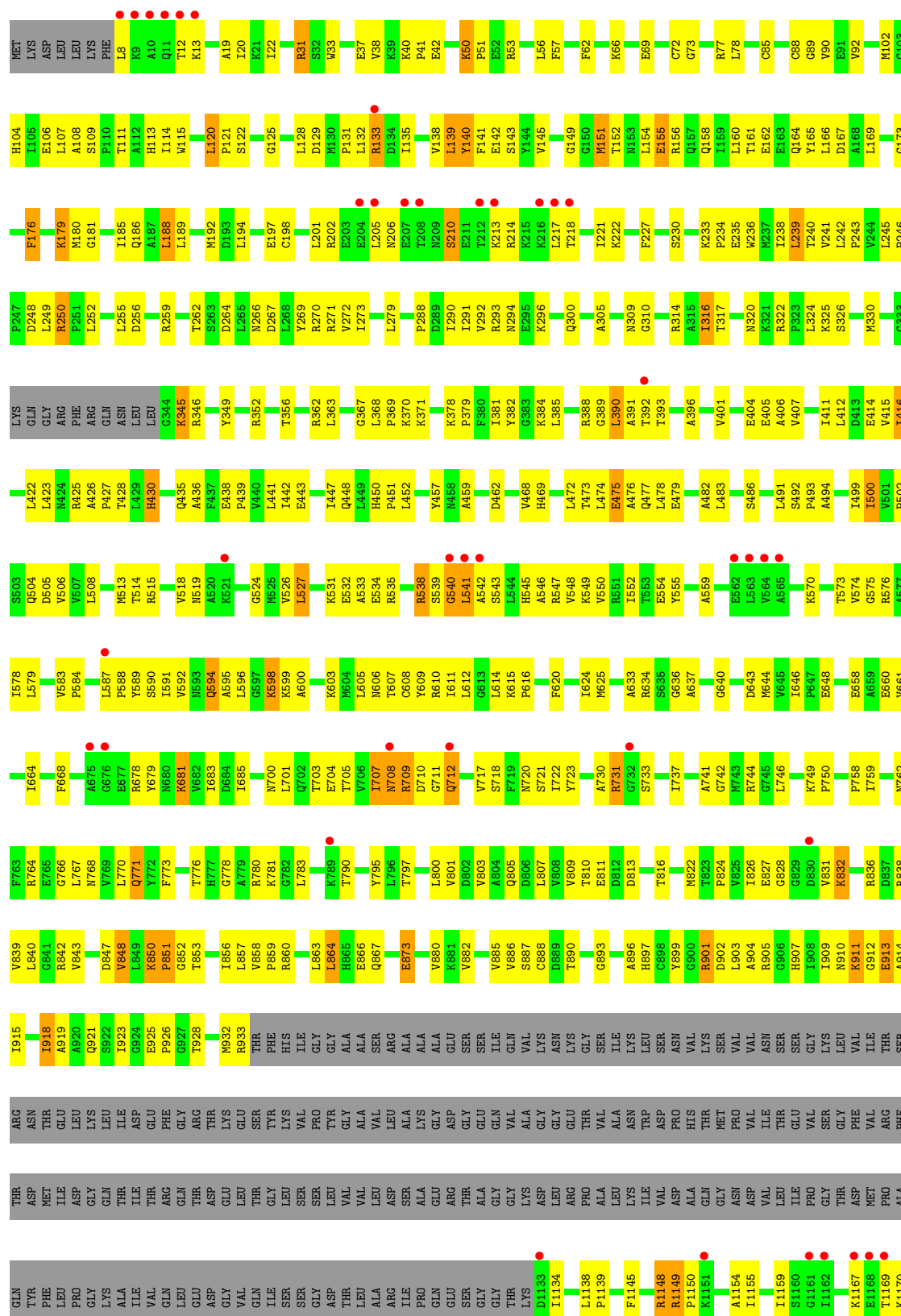


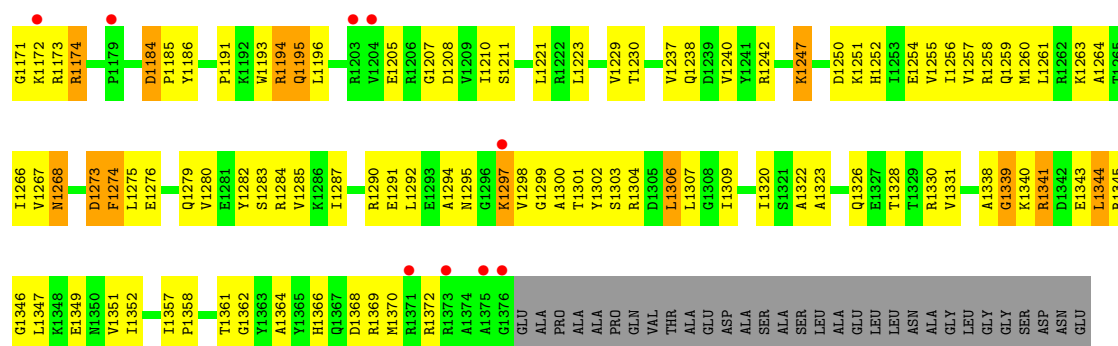


ALA	S1324	K1251	K1167	ASP	VAL	VAL	ASN	G906	G829	A741	A657	W580	1500	L422	G333	L239
GLU	F1325	H1252	E1168	VAL	ILE	THR	THR	H907	D830	G742	E658	V501	V501	L423	LYS	T240
LEU	Q1326	I1253	T1169	LEU	GLU	GLU	THR	I908	K832	W743	A659	V583	P502	H424	GLN	V241
ASN	I1327	V1254	L1170	PRO	VAL	VAL	GLY	I909	W831	R744	E660	P584	P503	R425	GLY	L242
ALA	T1328	V1255	K1171	GLY	THR	THR	LYS	N910	R836	G745	V661	K885	Q504	R426	ARG	P243
GLY	I1329	I1256	G1172	THR	GLY	GLY	LEU	G912	D837	L746	I664	G586	Q505	A426	PHE	L245
LEU	R1330	V1257	R1173	ASP	PHE	PHE	VAL	G913	R838	K749	F668	L587	V506	P427	ARG	P246
GLY	V1331	Q1258	R1174	MET	VAL	VAL	ILE	E914	W839	P750	F668	Y589	L508	T428	GLN	P247
GLY	A1338	M1260	L1175	PRO	ARG	ARG	THR	I915	L840	P758	L672	S590	R431	L429	ASN	D248
SER	G1339	R1261	V1176	ALA	PHE	PHE	THR	I916	G841	I759	L672	I591	R430	R429	LEU	L249
ASP	K1340	R1262	L1177	THR	THR	THR	ARG	I918	R842	P758	L672	I591	R430	R429	LEU	L249
ASN	R1341	K1263	T1178	TYR	ASP	ASP	ASN	I919	W843	I759	L672	I591	R430	R429	LEU	L249
GLU	D1342	A1264	P1179	PHE	MET	MET	THR	A920	R843	I759	L672	I591	R430	R429	LEU	L249
	I1343	T1265	P1179	LEU	ILE	ILE	THR	A921	R843	I759	L672	I591	R430	R429	LEU	L249
	L1344	V1266	D1184	PRO	ASP	ASP	GLU	Q921	D847	F763	R679	Q594	V518	I434	K345	L252
	R1345	V1267	P1185	GLY	THR	THR	LEU	S922	W848	F763	R679	Q594	V518	I434	K345	L252
	G1346	M1268	V1186	LYS	GLN	GLN	LYS	I923	L849	G765	K681	A595	N519	Q435	R346	L255
	L1347	A1269	E1187	ALA	THR	THR	ILE	E925	P851	G766	I683	G597	G522	F437	V347	D256
	K1348	E1188	E1188	ILE	THR	THR	ASP	P926	G852	L767	D684	K599	G524	E438	D348	Y349
	E1349	D1273	P1191	VAL	THR	THR	GLU	G927	I856	W768	I685	A600	N525	E443	T386	R259
	V1351	F1274	K1192	GLN	ARG	ARG	PHE	Q928	L857	A687	W686	K603	V526	E443	T386	F260
	I1352	Q1279	W1193	ASP	THR	THR	THR	L930	W858	L770	A688	K603	V526	E443	T386	F260
	I1357	V1280	R1194	ASP	GLU	GLU	LYS	M932	R860	H777	A689	M604	T528	I447	H364	Y269
	P1358	E1281	Q1195	VAL	GLY	GLY	THR	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	A1359	V1282	L1196	THR	LEU	LEU	GLU	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	G1360	S1283	M1197	THR	GLY	GLY	THR	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	T1361	R1284	V1198	ILE	GLY	GLY	THR	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	V1362	V1285	F1199	SER	LEU	LEU	THR	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	G1363	K1286	E1198	SER	SER	SER	VAL	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	A1364	I1287	R1203	GLY	THR	THR	PRO	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	V1365	V1290	V1204	ASP	LEU	LEU	THR	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	G1366	E1291	E1205	THR	VAL	VAL	GLY	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	D1367	L1292	D1208	ALA	ALA	ALA	VAL	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	R1368	V1209	V1209	ARG	ASP	ASP	LEU	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	R1369	I1210	I1210	ILE	SER	SER	ALA	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	M1370	S1211	S1211	PRO	ALA	ALA	GLY	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	R1371	V1298	V1298	GLN	GLY	GLY	THR	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	R1372	G1299	G1299	SER	THR	THR	GLY	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	G1376	A1300	G1225	GLY	ALA	ALA	GLU	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	GLU	T1301	V1226	THR	GLY	GLY	GLN	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	ALA	Y1302	H1227	THR	GLY	GLY	VAL	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	PRO	S1303	A1228	LYS	LYS	LYS	ALA	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	ALA	R1304	V1229	ASP	GLY	GLY	VAL	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	ALA	D1305	T1230	LEU	LEU	LEU	GLY	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	PRO	L1306	V1234	ARG	ARG	ARG	THR	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	GLN	L1307	V1234	THR	THR	THR	VAL	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	VAL	G1308	V1237	ALA	ALA	ALA	VAL	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	THR	I1309	Q1238	LYS	LYS	LYS	ALA	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	ALA	T1310	D1239	ILE	ILE	ILE	THR	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	GLU	K1311	V1240	VAL	VAL	VAL	THR	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	ASP	A1312	R1148	ASP	ASP	ASP	PRO	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	ALA	R1149	R1149	ASN	ASN	ASN	THR	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	SER	I1320	R1242	VAL	VAL	VAL	HIS	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	ALA	S1321	K1151	LYS	LYS	LYS	THR	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	SER	A1322	K1247	SER	GLY	GLY	MET	R933	R860	H777	A689	M604	T528	I447	H364	Y269
	LEU	A1323	K1247	ASN	ASN	ASN	PRO	R933	R860	H777	A689	M604	T528	I447	H364	Y269

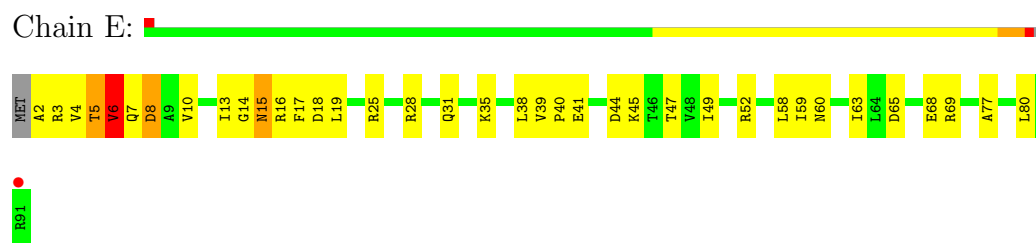
## ● Molecule 3: Escherichia coli RNA polymerase beta' subunit

Chain I:

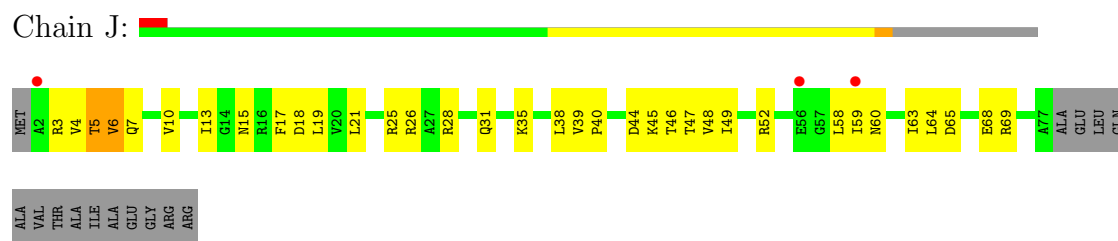




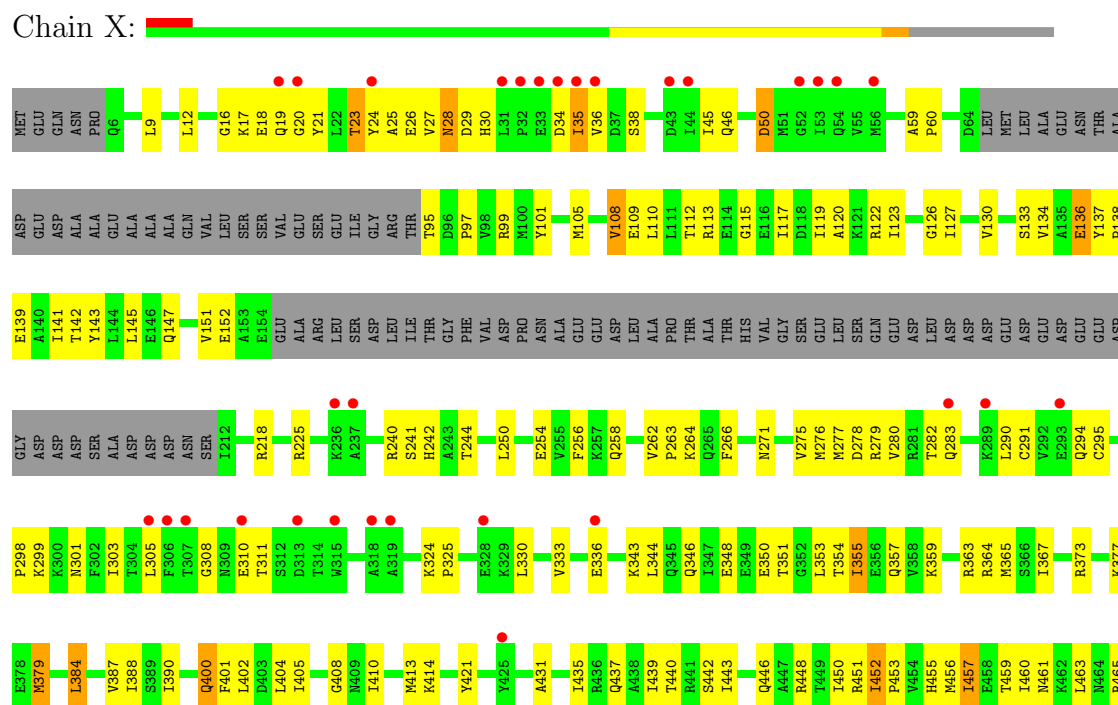
• Molecule 4: Escherichia coli RNA polymerase omega subunit

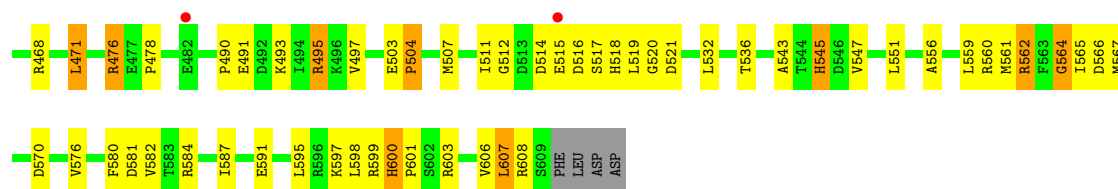


• Molecule 4: Escherichia coli RNA polymerase omega subunit



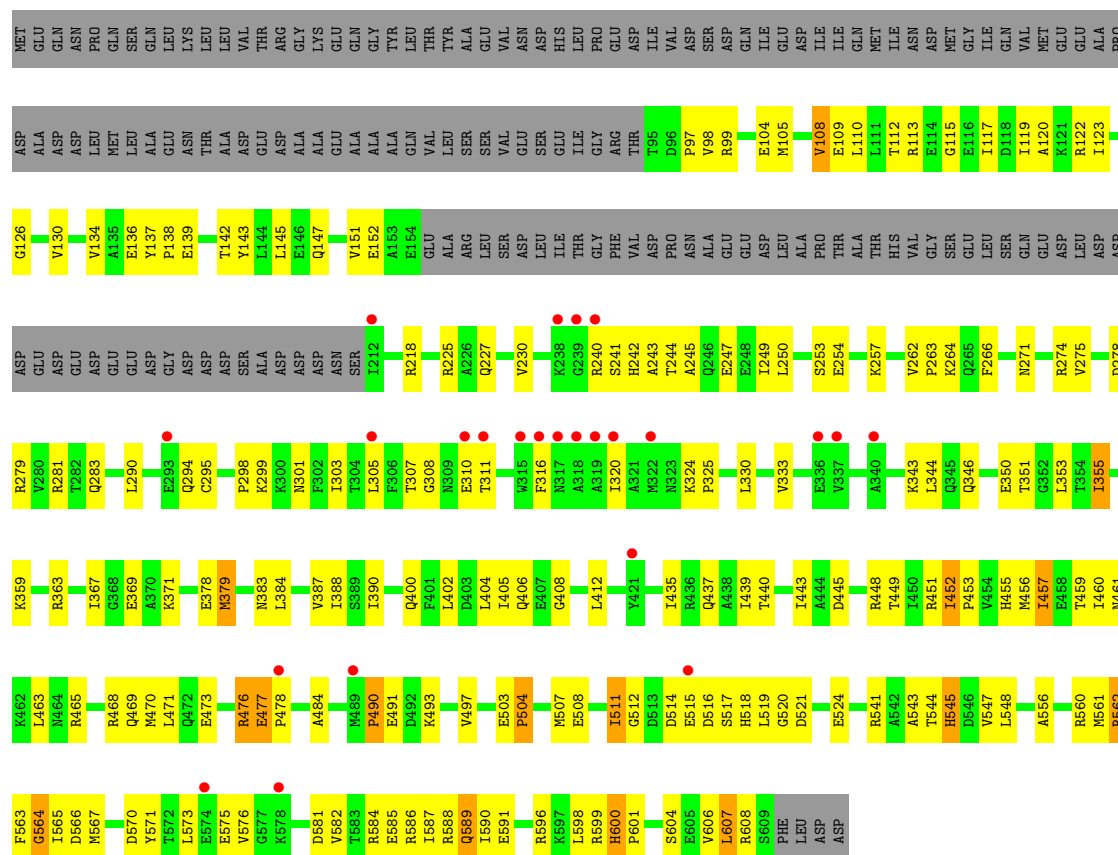
• Molecule 5: Escherichia coli RNA polymerase sigma70 subunit





• Molecule 5: Escherichia coli RNA polymerase sigma70 subunit

Chain Y:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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) 82.3 (29.84-3.85)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.33 (at 3.86Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1_1168)	Depositor
R, $R_{free}$	0.252 , 0.320 0.257 , 0.323	Depositor DCC
$R_{free}$ test set	4498 reflections (4.98%)	DCC
Wilson B-factor (Å <sup>2</sup> )	122.2	Xtriage
Anisotropy	0.142	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , -6.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.40$ , $\langle L^2 \rangle = 0.22$	Xtriage
Outliers	0 of 96273 reflections	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	56126	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.71% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2514	0	2566	97	0
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	0	2	0
All	All	56115	11	56806	2538	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 22.

The worst 5 of 2538 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(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

### 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	A	321/329 (98%)	266 (83%)	41 (13%)	14 (4%)	4 47

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	217/329 (66%)	186 (86%)	24 (11%)	7 (3%)	6	57
1	F	227/329 (69%)	196 (86%)	26 (12%)	5 (2%)	10	65
1	G	213/329 (65%)	188 (88%)	21 (10%)	4 (2%)	12	69
2	C	1333/1342 (99%)	1069 (80%)	213 (16%)	51 (4%)	5	52
2	H	1333/1342 (99%)	1070 (80%)	213 (16%)	50 (4%)	5	52
3	D	1154/1407 (82%)	922 (80%)	189 (16%)	43 (4%)	5	53
3	I	1154/1407 (82%)	929 (80%)	183 (16%)	42 (4%)	5	54
4	E	88/91 (97%)	76 (86%)	8 (9%)	4 (4%)	4	46
4	J	74/91 (81%)	64 (86%)	6 (8%)	4 (5%)	3	41
5	X	511/613 (83%)	450 (88%)	45 (9%)	16 (3%)	7	58
5	Y	454/613 (74%)	411 (90%)	32 (7%)	11 (2%)	9	64
All	All	7079/8222 (86%)	5827 (82%)	1001 (14%)	251 (4%)	6	55

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%)	60	91
1	B	189/286 (66%)	185 (98%)	4 (2%)	66	92
1	F	197/286 (69%)	192 (98%)	5 (2%)	60	91
1	G	185/286 (65%)	180 (97%)	5 (3%)	57	90
2	C	1150/1157 (99%)	1087 (94%)	63 (6%)	30	78
2	H	1150/1157 (99%)	1092 (95%)	58 (5%)	34	80

Continued on next page...



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	D	971/1168 (83%)	919 (95%)	52 (5%)	31	78
3	I	971/1168 (83%)	921 (95%)	50 (5%)	33	80
4	E	74/75 (99%)	71 (96%)	3 (4%)	41	84
4	J	65/75 (87%)	65 (100%)	0	100	100
5	X	460/540 (85%)	442 (96%)	18 (4%)	43	85
5	Y	407/540 (75%)	391 (96%)	16 (4%)	43	85
All	All	6100/7024 (87%)	5819 (95%)	281 (5%)	37	82

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 chains 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 ⓘ

There are no carbohydrates 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
7	G4P	D	1503	-	38,38,38	2.31	11 (28%)	58,61,61	7.45	12 (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	-	-	0/27/43/43	0/1/3/3

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	D	1503	G4P	C2-N2	6.34	1.42	1.32
7	D	1503	G4P	C2'-C1'	-5.39	1.45	1.53
7	D	1503	G4P	C2-N3	5.06	1.39	1.33
7	D	1503	G4P	C5-C4	-3.82	1.31	1.40
7	D	1503	G4P	C8-N9	3.64	1.42	1.36

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	D	1503	G4P	C6-C5-N7	-54.91	126.75	134.14
7	D	1503	G4P	C4'-O4'-C1'	-6.45	102.75	109.75
7	D	1503	G4P	O4'-C1'-N9	5.64	113.69	108.44
7	D	1503	G4P	PC-O3C-PD	-4.95	117.16	131.68
7	D	1503	G4P	O4'-C1'-C2'	-4.36	100.09	106.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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 ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	A	323/329 (98%)	0.25	9 (2%)	50	38	1, 73, 178, 299	0
1	B	221/329 (67%)	0.43	12 (5%)	25	20	5, 98, 204, 259	0
1	F	229/329 (69%)	0.47	11 (4%)	29	24	27, 123, 199, 278	0
1	G	217/329 (65%)	0.46	11 (5%)	27	22	34, 118, 187, 236	0
2	C	1335/1342 (99%)	0.11	35 (2%)	53	40	0, 48, 170, 262	0
2	H	1335/1342 (99%)	0.29	50 (3%)	39	31	0, 88, 203, 293	0
3	D	1160/1407 (82%)	0.18	38 (3%)	44	34	0, 42, 158, 289	0
3	I	1160/1407 (82%)	0.29	49 (4%)	35	27	0, 60, 192, 317	0
4	E	90/91 (98%)	0.05	1 (1%)	77	61	1, 52, 121, 158	0
4	J	76/91 (83%)	0.43	3 (3%)	37	29	10, 89, 165, 211	0
5	X	517/613 (84%)	0.31	33 (6%)	19	17	0, 101, 226, 326	0
5	Y	458/613 (74%)	0.33	24 (5%)	26	22	1, 109, 234, 357	0
All	All	7121/8222 (86%)	0.25	276 (3%)	37	29	0, 73, 197, 357	0

The worst 5 of 276 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	10	ALA	10.7
3	I	11	GLN	7.7
2	H	1000	LEU	6.4
5	X	35	ILE	6.4
5	X	36	VAL	5.6

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

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

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
7	G4P	D	1503	36/36	0.23	-0.38	31,56,93,118	0
6	ZN	D	1502	1/1	0.18	-0.63	8,8,8,8	0
6	ZN	I	1502	1/1	0.15	-0.78	49,49,49,49	0
6	ZN	D	1501	1/1	0.08	-1.45	54,54,54,54	0
6	ZN	I	1501	1/1	0.05	-1.85	60,60,60,60	0

### 6.5 Other polymers ⓘ

There are no such residues in this entry.