



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:51 PM GMT

PDB ID : 4JK2
Title : X-ray crystal structure of Escherichia coli sigma70 holoenzyme in complex with
guanosine pentaphosphate (pppGpp)
Authors : Murakami, K.S.
Deposited on : 2013-03-09
Resolution : 4.20 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
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

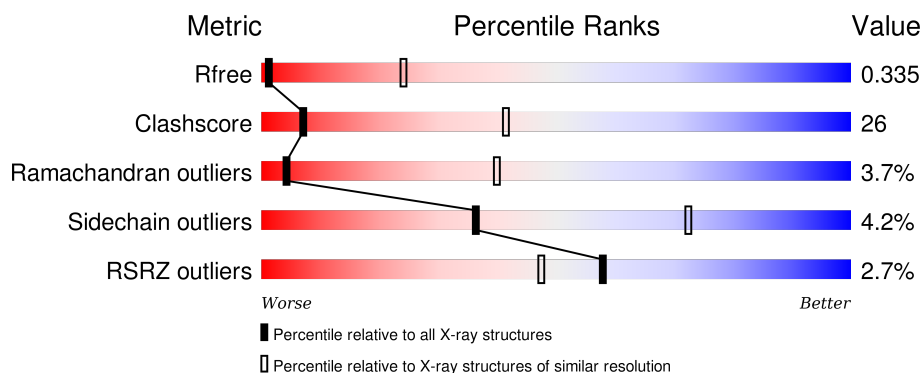
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.20 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}	91344	1039 (4.80-3.60)
Clashscore	102246	1140 (4.80-3.60)
Ramachandran outliers	100387	1083 (4.80-3.60)
Sidechain outliers	100360	1067 (4.80-3.60)
RSRZ outliers	91569	1042 (4.80-3.60)

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. 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>2%</div> <div>59% 36% . .</div> </div>
1	B	329	<div> <div>2%</div> <div>38% 27% . 33%</div> </div>
1	F	329	<div> <div>4%</div> <div>46% 22% . 30%</div> </div>
1	G	329	<div> <div>2%</div> <div>40% 24% . 34%</div> </div>
2	C	1342	<div> <div>2%</div> <div>53% 41% 5% .</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	H	1342	
3	D	1407	
3	I	1407	
4	E	91	
4	J	91	
5	X	613	
5	Y	613	

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
7	0O2	D	1503	-	-	X	-

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 56129 atoms, of which 10 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			

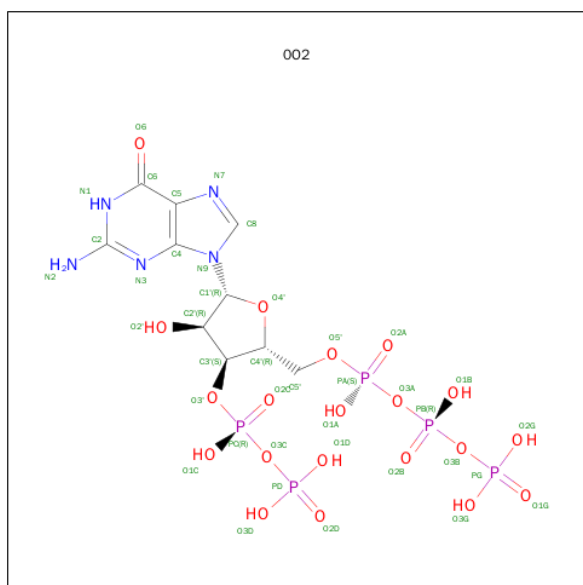
- Molecule 5 is a protein called Escherichia coli RNA polymerase sigma70 subunit.

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		

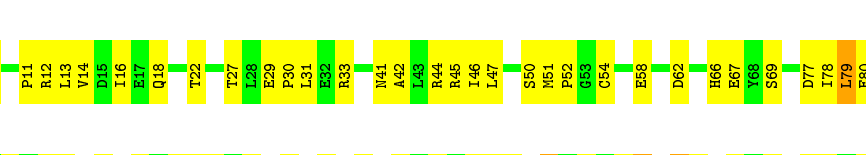
- Molecule 7 is GUANOSINE 5'-(TETRAHYDROGEN TRIPHOSPHATE) 3'-(TRIHYDROGEN DIPHOSPHATE) (three-letter code: 002) (formula: C₁₀H₁₈N₅O₂₀P₅).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	D	1	Total	C	H	N	O	0	0
			50	10	10	5	20		

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

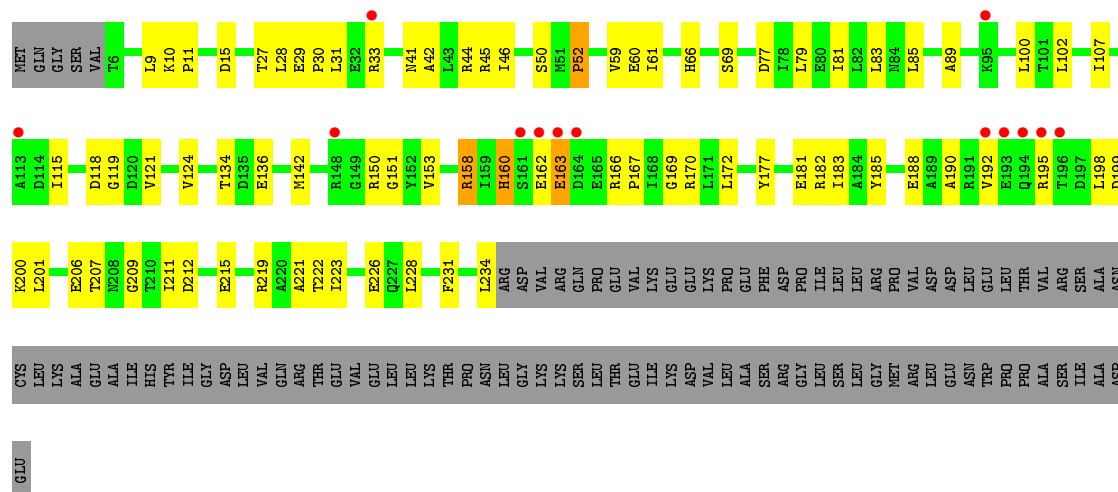
- Chain A: 



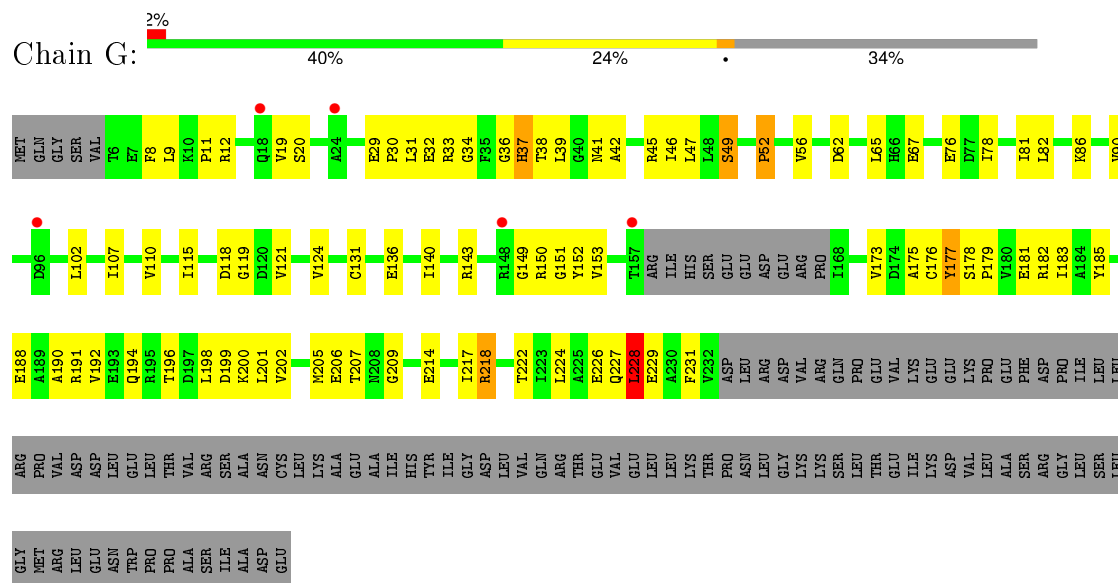
Position	Amino Acid	Frequency (bits)
1	MET	0.15
2	GLN	0.15
3	G3	0.15
4	S4	0.15
5	F8	0.15
6	P11	0.15
7	R12	0.15
8	L13	0.15
9	V14	0.15
10	D15	0.15
11	I16	0.15
12	E17	0.15
13	Q18	0.15
14	T22	0.15
15	T27	0.15
16	L28	0.15
17	E29	0.15
18	P30	0.15
19	L31	0.15
20	E32	0.15
21	R33	0.15
22	N41	0.15
23	A42	0.15
24	L43	0.15
25	R44	0.15
26	R45	0.15
27	L46	0.15
28	L47	0.15
29	S50	0.15
30	M51	0.15
31	P52	0.15
32	G53	0.15
33	C54	0.15
34	E58	0.15
35	D62	0.15
36	H66	0.15
37	E67	0.15
38	Y68	0.15
39	S69	0.15
40	D77	0.15
41	L78	0.15
42	L79	0.15
43	E80	0.15
44	L83	0.15
45	K86	0.15
46	G87	0.15
47	L88	0.15
48	L89	0.15
49	L90	0.15
50	V90	0.15
51	Q93	0.15
52	L100	0.15
53	T101	0.15
54	L102	0.15
55	N103	0.15
56	K104	0.15
57	S105	0.15
58	V110	0.15
59	I115	0.15
60	H117	0.15
61	D118	0.15
62	G119	0.15
63	D120	0.15
64	V121	0.15
65	G131	0.15
66	T134	0.15
67	R143	0.15
68	R148	0.15
69	G149	0.15
70	R150	0.15
71	G151	0.15
72	Y152	0.15
73	V153	0.15
74	P154	0.15
75	R158	0.15
76	L159	0.15
77	H160	0.15
78	S161	0.15
79	E162	0.15
80	E163	0.15
81	R166	0.15
82	P167	0.15
83	R170	0.15
84	L171	0.15
85	L172	0.15
86	S178	0.15
87	P179	0.15
88	V180	0.15
89	E181	0.15
90	R182	0.15
91	I183	0.15
92	V187	0.15
93	E188	0.15
94	L189	0.15
95	A190	0.15

- Chain B:
-
- 38% 27% 33%
- 2%
- MET GLN GLY SER VAL T6 E7 F8 L9 K10 P11 R12 L13 I16 E17 Q18 V19 S20 S21 T22 V26 T27 L28 E29 P30 L31 E32 R33 H37 T38 L39 S40 M41 A42 R45 L46 Y47 L48 S50 S49 M51 P52 T61 D62 G63 V64 L65 S69 G73 V74 Q75 E76 D77

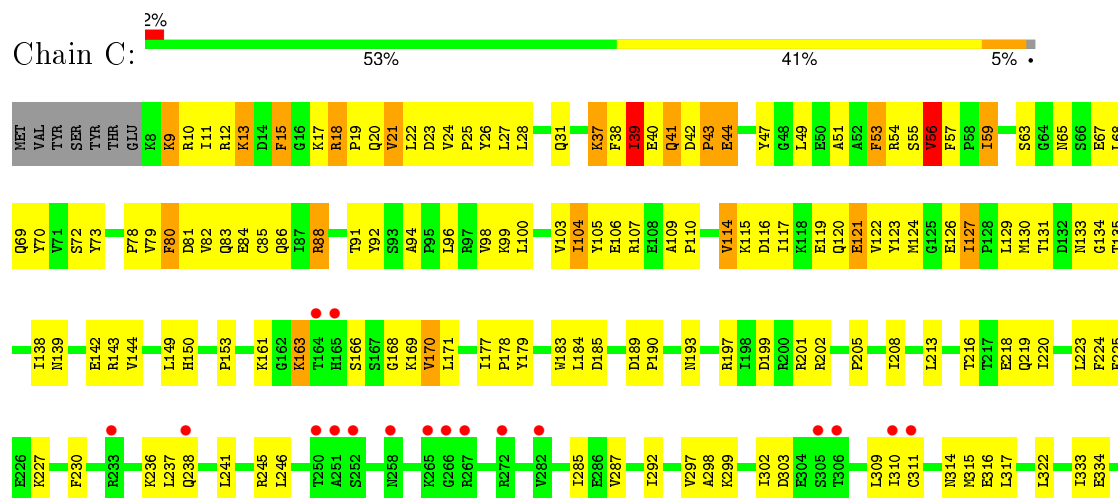
- Chain F: 



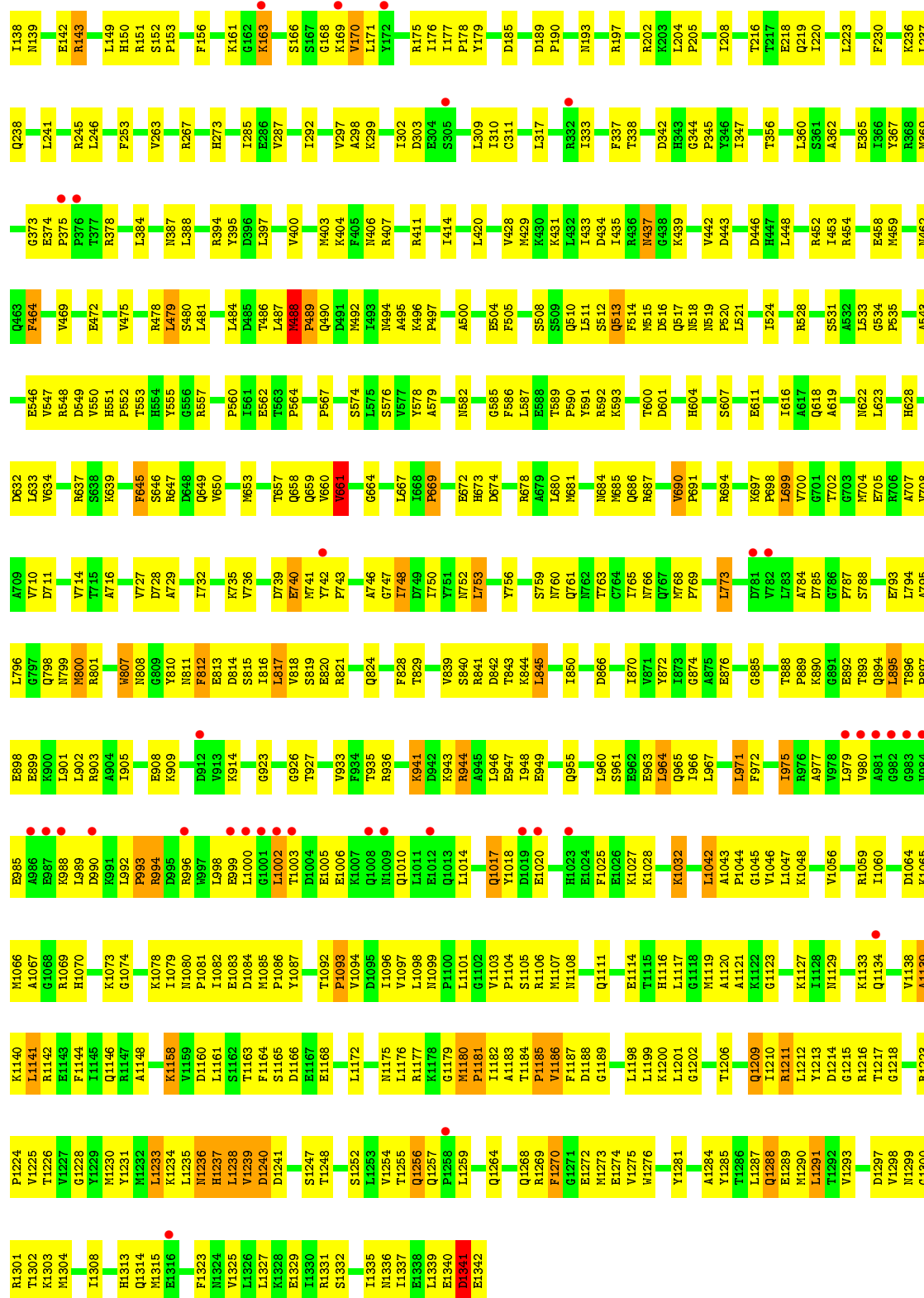
- Molecule 1: Escherichia coli RNA polymerase alpha subunit



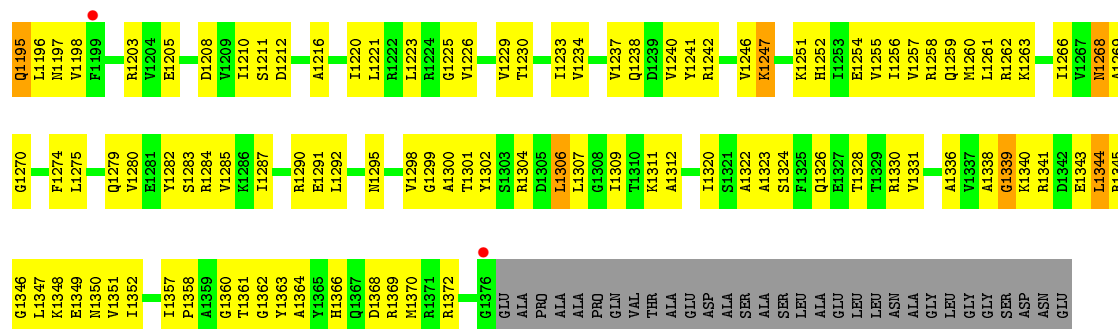
- Molecule 2: Escherichia coli RNA polymerase beta subunit



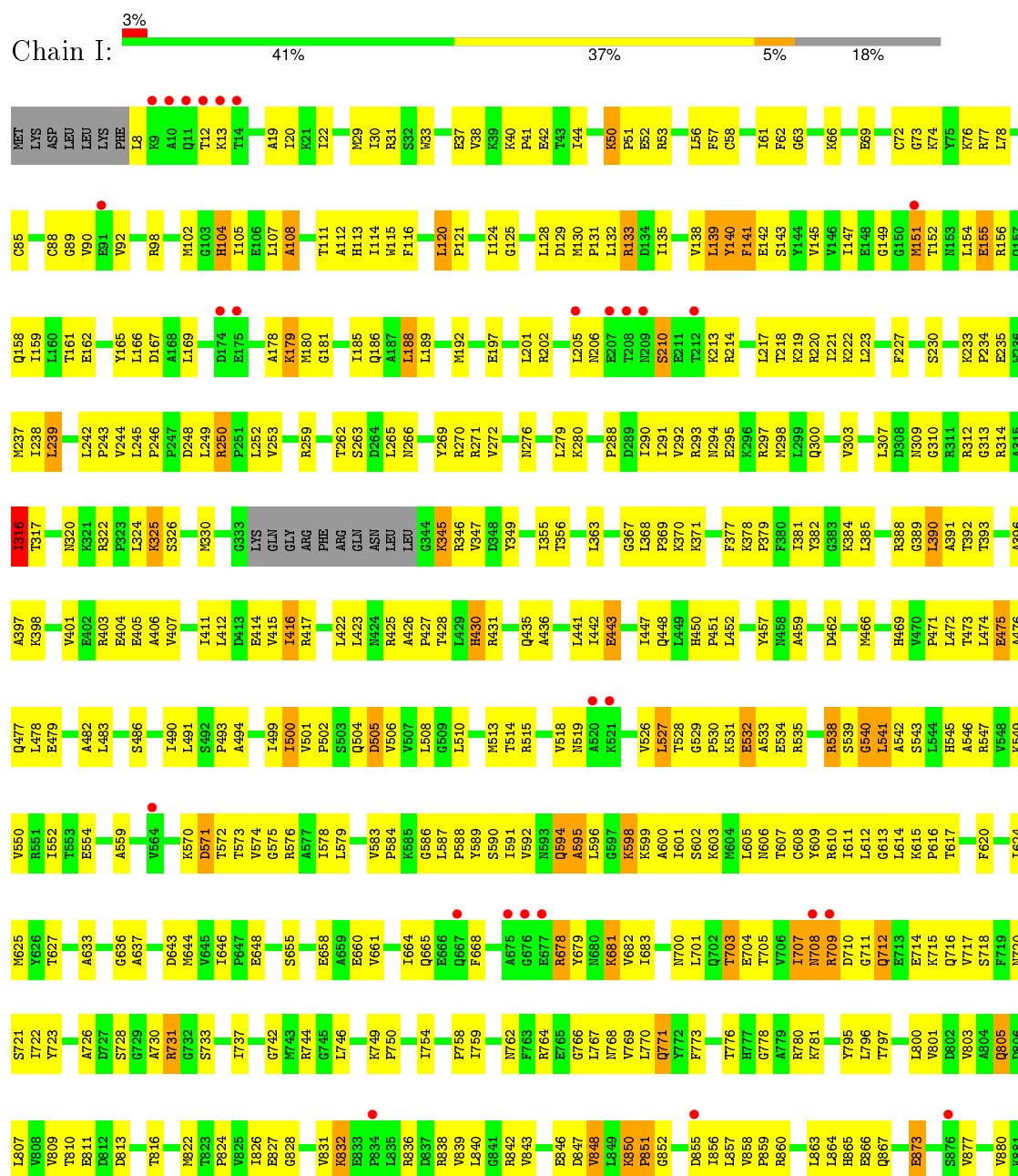


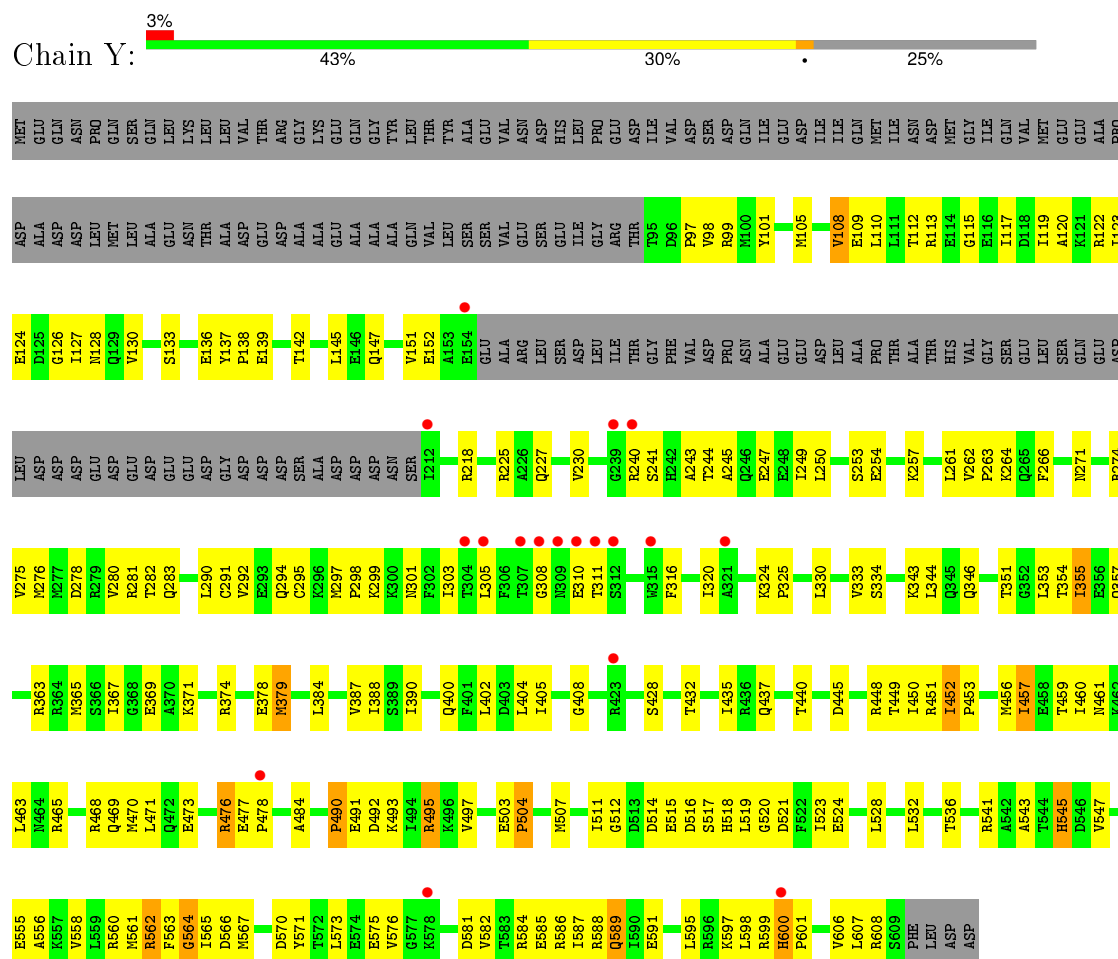


GLY	ASP	THR	LEU	SER	PRO	GLY	L871	D802	S718	N625	I552	L474	T393	I309	L217	V146	K79	MET
THR	LEU	GLY	VAL	LEU	TYR	GLY	L872	V803	F719	N626	I553	E475	I394	G310	L218	I147	H80	LYS
LEU	VAL	ALA	VAL	ASP	GLY	ALA	E873	Q805	N720	T627	E554	A476	K395	R311	K219	E148	H81	ASP
ALA	VAL	ALA	VAL	VAL	VAL	SER	V877	D806	I722	Q628	Y555	L478	A396	R312	R220	G149	LEU	LEU
SER	ASP	ALA	SER	ALA	ALA	ARG	V878	N807	I723	A633	D558	A482	K397	G313	K222	M151	C85	PHE
ALA	SER	ALA	ALA	ALA	ALA	ALA	V879	V808	N724	D643	A559	L483	A398	R314	T152	T152	PHE	LYS
PRO	GLY	ALA	GLY	GLY	GLY	ALA	V880	T810	N725	N644	D571	L483	E404	T317	S230	H153	C88	LYS
GLY	GLY	ALA	GLY	GLY	GLY	ALA	R881	T811	S728	N645	T572	T487	E405	T317	K233	H154	G89	LYS
ARG	ARG	GLY	GLY	GLY	GLY	SER	R882	E811	N729	N646	T573	T487	E406	T317	P234	E155	V90	LYS
GLY	GLY	ALA	GLY	GLY	GLY	SER	R883	D812	N730	N647	T574	T487	A407	T317	E235	E156	E91	LYS
GLY	GLY	ALA	GLY	GLY	GLY	SER	S884	D813	N731	N648	G575	T487	A408	T317	E236	Q157	V92	LYS
THR	GLY	VAL	GLY	VAL	VAL	GLN	V885	T816	N732	N649	R576	L491	V408	T317	E237	I159	Q94	LYS
LYS	LYS	LYS	LYS	LYS	LYS	VAL	V886	H817	S733	N653	A577	P498	N409	G333	L238	L160	R98	LYS
D1133	D1133	D1134	D1134	D1134	D1134	LYS	C888	E818	I737	N654	I578	P499	N410	G334	L239	E162	R99	LYS
L1138	L1138	L1139	L1139	L1139	L1139	LYS	D899	N820	I737	N655	L579	I500	D413	G335	L240	E162	R101	LYS
P1139	P1139	P1140	P1140	P1140	P1140	GLY	T890	T823	A741	N656	N820	V501	P243	R102	P243	L166	M102	LYS
F1145	F1145	F1146	F1146	F1146	F1146	GLY	G893	P824	G742	N657	V833	P502	V244	R103	P244	D167	M103	LYS
R1148	R1148	R1149	R1149	R1149	R1149	LYS	A896	V825	N743	N658	P884	S503	L245	R104	P245	A168	G103	LYS
R1149	R1149	R1150	R1150	R1150	R1150	LYS	B897	I826	N744	N659	G885	Q504	R417	R105	P246	L169	H104	LYS
P1150	P1150	P1151	P1151	P1151	P1151	LEU	G898	E827	N745	N660	G886	D505	P247	R106	P247	G104	I105	LYS
E1158	E1158	E1159	E1159	E1159	E1159	ASP	C899	G828	N746	N661	L587	V506	D248	R107	P248	G105	E106	LYS
R1163	R1163	R1164	R1164	R1164	R1164	ASN	Y899	G829	N747	N662	P888	V507	L249	R108	P249	D174	L107	LYS
R1164	R1164	R1165	R1165	R1165	R1165	VAL	G900	V831	N748	N663	Y899	L508	R250	R109	P250	E175	A108	LYS
E1168	E1168	E1169	E1169	E1169	E1169	THR	R901	R832	N749	N664	Y899	L509	R251	R110	P251	F176	T111	LYS
R1170	R1170	R1171	R1171	R1171	R1171	LYS	D902	E833	N750	N665	S590	M513	L252	R111	P252	D177	A112	LYS
R1171	R1171	R1172	R1172	R1172	R1172	VAL	L903	P834	N751	N666	I591	T514	P427	R112	P253	A178	H113	LYS
R1172	R1172	R1173	R1173	R1173	R1173	VAL	A904	N835	N752	N667	V592	R515	T428	R113	P254	K179	H114	LYS
R1173	R1173	R1174	R1174	R1174	R1174	ASN	R905	G836	N753	N668	G594	R516	T429	R114	P255	M180	T114	LYS
E1175	E1175	E1176	E1176	E1176	E1176	SER	G906	R837	N754	N669	A595	V518	R430	R115	P256	G181	W115	LYS
T1176	T1176	T1177	T1177	T1177	T1177	GLY	R907	D838	N755	N670	A596	N519	R431	R116	P257	I185	K118	LYS
R1177	R1177	R1178	R1178	R1178	R1178	GLY	R908	V839	N756	N671	L596	A520	A436	R117	P258	Q186	S119	LYS
V1180	V1180	V1181	V1181	V1181	V1181	GLY	R909	L840	N757	N672	G597	K521	A437	R118	P259	A187	P121	LYS
D1184	D1184	D1185	D1185	D1185	D1185	GLY	N910	G841	N758	N673	K598	G522	A438	R119	P260	L188	P122	LYS
P1185	P1185	P1186	P1186	P1186	P1186	VAL	K911	R842	N759	N674	K599	G523	A439	R120	P261	L189	P123	LYS
E1187	E1187	E1188	E1188	E1188	E1188	VAL	G912	V943	N760	N675	A600	G524	A440	R121	P262	E190	P124	LYS
P1191	P1191	P1192	P1192	P1192	P1192	THR	E913	D847	N761	N676	S802	H525	A441	R122	P263	M192	G125	LYS
K1193	K1193	K1194	K1194	K1194	K1194	THR	A914	V848	N762	N677	G603	V526	A442	R123	P264	D193	L126	LYS
R1194	R1194	R1195	R1195	R1195	R1195	THR	T915	L849	N763	N678	N604	L527	A443	R124	P265	L194	L127	LYS
E1198	E1198	E1199	E1199	E1199	E1199	THR	R198	R850	N764	N679	L605	K531	A444	R125	P266	E197	D128	LYS
P1199	P1199	P1200	P1200	P1200	P1200	THR	A919	P851	N765	N680	T807	E532	A445	R126	P267	L201	D129	LYS
E1203	E1203	E1204	E1204	E1204	E1204	THR	G921	G852	N766	N681	C608	E533	A446	R127	P268	L202	P131	LYS
D1204	D1204	D1205	D1205	D1205	D1205	THR	L930	D855	N767	N682	R610	R535	A447	R128	P269	R202	G63	LYS
P1205	P1205	P1206	P1206	P1206	P1206	THR	E925	L857	N768	N683	I611	R536	A448	R129	P270	L203	G64	LYS
E1206	E1206	E1207	E1207	E1207	E1207	THR	P926	V858	N769	N684	I612	R537	A449	R130	P271	L204	G65	LYS
P1207	P1207	P1208	P1208	P1208	P1208	THR	G927	R860	N770	N685	G613	L541	A450	R131	P272	L205	G66	LYS
E1208	E1208	E1209	E1209	E1209	E1209	THR	L931	N861	N771	N686	L542	A542	A451	R132	P273	L206	G67	LYS
P1209	P1209	P1210	P1210	P1210	P1210	THR	E928	V862	N772	N687	R615	R543	A452	R133	P274	L207	G68	LYS
E1210	E1210	E1211	E1211	E1211	E1211	THR	G929	R863	N773	N688	I616	R544	A453	R134	P275	L208	G69	LYS
P1211	P1211	P1212	P1212	P1212	P1212	THR	R930	N863	N774	N689	T817	R545	A454	R135	P276	L209	G70	LYS
E1212	E1212	E1213	E1213	E1213	E1213	THR	R931	N864	N775	N690	T818	R546	A455	R136	P277	L210	G71	LYS
P1213	P1213	P1214	P1214	P1214	P1214	THR	R932	N865	N776	N691	T819	R547	A456	R137	P278	L211	G72	LYS
E1214	E1214	E1215	E1215	E1215	E1215	THR	R933	N866	N777	N692	T820	R548	A457	R138	P279	L212	G73	LYS
P1215	P1215	P1216	P1216	P1216	P1216	THR	R934	N867	N778	N693	T821	R549	A458	R139	P280	L213	G74	LYS
E1216	E1216	E1217	E1217	E1217	E1217	THR	R935	N868	N779	N694	T822	R550	A459	R140	P281	L214	G75	LYS
P1217	P1217	P1218	P1218	P1218	P1218	THR	R936	N869	N780	N695	T823	R551	A460	R141	P282	L215	G76	LYS
E1218	E1218	E1219	E1219	E1219	E1219	THR	R937	N870	N781	N696	T824	R552	A461	R142	P283	L216	G77	LYS
P1219	P1219	P1220	P1220	P1220	P1220	THR	R938	N871	N782	N697	T825	R553	A462	R143	P284	L217	G78	LYS
E1220	E1220	E1221	E1221	E1221	E1221	THR	R939	N872	N783	N698	T826	R554	A463	R144	P285	L218	G79	LYS
P1221	P1221	P1222	P1222	P1222	P1222	THR	R940	N873	N784	N699	T827	R555	A464	R145	P286	L219	G80	LYS
E1222	E1222	E1223	E1223	E1223	E1223	THR	R941	N874	N785	N700	T828	R556	A465	R146	P287	L220	G81	LYS
P1223	P1223	P1224	P1224	P1224	P1224	THR	R942	N875	N786	N701	T829	R557	A466	R147	P288	L221	G82	LYS
E1224	E1224	E1225	E1225	E1225	E1225	THR	R943	N876	N787	N702	T830	R558	A467	R148	P289	L222	G83	LYS
P1225	P1225	P1226	P1226	P1226	P1226	THR	R944	N877	N788	N703	T831	R559	A468	R149	P290	L223	G84	LYS
E1226	E1226	E1227	E1227	E1227	E1227	THR	R945	N878	N789	N704	T832	R560	A469	R150	P291	L224	G85	LYS
P1227	P1227	P1228	P1228	P1228	P1228	THR	R946	N879	N790	N705	T833	R561	A470	R151	P292	L225	G86	LYS
E1228	E1228	E1229	E1229	E1229	E1229	THR	R947	N880	N791	N706	T834	R562	A471	R152	P293	L226	G87	LYS
P1229	P1229	P1230	P1230	P1230	P1230	THR	R948	N881	N792	N707	T835	R563	A472	R153	P294	L227	G88	LYS
E1230	E1230	E1231	E1231	E1231	E1231	THR	R949	N882	N793	N708	T836	R564	A473	R154	P295	L228	G89	LYS
P1231	P1231	P1232	P1232	P1232	P1232	THR	R950	N883	N794	N709	T837	R565	A474	R155	P296	L229	G90	LYS
E1232	E1232	E1233	E1233	E1233	E1233	THR	R951	N884	N795	N710	T838	R566	A475	R156	P297	L230	G91	LYS
P1233	P1233	P1234	P1234	P1234	P1234	THR	R952	N885	N796	N711	T839	R567	A476	R157	P298	L231	G92	LYS
E1234	E1234	E1235	E1235	E1235	E1235	THR	R953	N886	N797	N712	T840	R568	A477	R158	P299	L232	G93	LYS
P1235	P1235	P1236	P1236	P1236	P1236	THR	R954	N887	N798	N713	T841	R569	A478	R159	P300	L233	G94	LYS
E1236	E1236	E1237	E1237	E1237	E1237	THR	R955	N888	N799	N714	T842	R570	A479	R160	P301	L234	G95	LYS
P1237	P1237	P1238	P1238	P1238	P1238	THR	R956	N889	N800	N715	T843	R571	A480	R161	P302	L235	G96	LYS
E1238	E1238	E1239	E1239	E1239	E1239	THR	R957	N890	N801	N716	T844	R572	A481	R162	P303	L236	G97	LYS
P1239	P1239	P1240	P1240	P1240	P1240	THR	R958	N891	N802	N717	T845	R573	A482	R163	P304	L237	G98	LYS
E1240	E1240	E1241	E1241	E1241	E1241	THR	R959	N892</										



● Molecule 3: Escherichia coli RNA polymerase beta' subunit





4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	185.32Å 205.41Å 309.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.94 – 4.20 30.04 – 4.20	Depositor EDS
% Data completeness (in resolution range)	80.4 (29.94-4.20) 70.4 (30.04-4.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.11 (at 4.26Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.244 , 0.322 0.264 , 0.335	Depositor DCC
R_{free} test set	3080 reflections (5.06%)	DCC
Wilson B-factor (Å ²)	159.7	Xtriage
Anisotropy	0.159	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 57.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	0 of 69537 reflections	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	56129	wwPDB-VP
Average B, all atoms (Å ²)	76.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.39% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: ZN, 0O2

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	0.36	0/3454
1	B	0.19	0/1725	0.39	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	0.38	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	0.19	0/4253	0.36	0/5719
5	Y	0.20	0/3783	0.36	0/5083
All	All	0.20	0/56889	0.38	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	110	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	91	0
1	F	1775	0	1800	79	0
1	G	1671	0	1706	91	0
2	C	10523	0	10546	600	0
2	H	10523	0	10546	574	0
3	D	9060	0	9257	658	0
3	I	9060	0	9257	591	0
4	E	708	0	719	52	0
4	J	605	0	612	33	0
5	X	4198	0	4250	197	0
5	Y	3732	0	3809	157	0
6	D	2	0	0	0	0
6	I	2	0	0	0	0
7	D	40	10	16	9	0
All	All	56119	10	56822	2973	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 26.

The worst 5 of 2973 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	1.23	1.20
3:D:1173:ARG:HA	3:D:1174:ARG:HB2	1.24	1.18
2:H:488:MET:HB2	2:H:490:GLN:H	1.07	1.11
3:I:850:LYS:HD2	3:I:851:PRO:HD2	1.30	1.09
2:H:1073:LYS:HD3	3:I:462:ASP:HB3	1.28	1.08

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%)	40 (12%)	15 (5%)	3	33
1	B	217/329 (66%)	187 (86%)	22 (10%)	8 (4%)	4	40
1	F	227/329 (69%)	193 (85%)	28 (12%)	6 (3%)	7	47
1	G	213/329 (65%)	186 (87%)	22 (10%)	5 (2%)	8	50
2	C	1333/1342 (99%)	1073 (80%)	208 (16%)	52 (4%)	4	37
2	H	1333/1342 (99%)	1078 (81%)	206 (16%)	49 (4%)	4	40
3	D	1154/1407 (82%)	926 (80%)	182 (16%)	46 (4%)	4	37
3	I	1154/1407 (82%)	925 (80%)	184 (16%)	45 (4%)	4	37
4	E	88/91 (97%)	77 (88%)	6 (7%)	5 (6%)	2	28
4	J	74/91 (81%)	64 (86%)	6 (8%)	4 (5%)	2	30
5	X	511/613 (83%)	450 (88%)	46 (9%)	15 (3%)	6	45
5	Y	454/613 (74%)	409 (90%)	34 (8%)	11 (2%)	7	49
All	All	7079/8222 (86%)	5834 (82%)	984 (14%)	261 (4%)	4	40

5 of 261 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%)	273 (97%)	8 (3%)	51	80
1	B	189/286 (66%)	186 (98%)	3 (2%)	70	88
1	F	197/286 (69%)	194 (98%)	3 (2%)	72	89
1	G	185/286 (65%)	182 (98%)	3 (2%)	70	88

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	C	1150/1157 (99%)	1094 (95%)	56 (5%)	31	69
2	H	1150/1157 (99%)	1097 (95%)	53 (5%)	33	70
3	D	971/1168 (83%)	921 (95%)	50 (5%)	29	68
3	I	971/1168 (83%)	918 (94%)	53 (6%)	27	66
4	E	74/75 (99%)	72 (97%)	2 (3%)	52	80
4	J	65/75 (87%)	65 (100%)	0	100	100
5	X	460/540 (85%)	447 (97%)	13 (3%)	51	80
5	Y	407/540 (75%)	392 (96%)	15 (4%)	41	75
All	All	6100/7024 (87%)	5841 (96%)	259 (4%)	36	72

5 of 259 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	E	8	ASP
2	H	56	VAL
3	I	1149	ARG
5	X	28	ASN
5	X	607	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 91 such sidechains are listed below:

Mol	Chain	Res	Type
5	X	54	GLN
1	G	37	HIS
5	Y	301	ASN
5	X	258	GLN
5	X	437	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	0O2	D	1503	-	31,42,42	2.05	9 (29%)	46,68,68	2.32	10 (21%)

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	0O2	D	1503	-	-	0/29/49/49	0/3/3/3

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
7	D	1503	0O2	O2'-C2'	-4.34	1.32	1.43
7	D	1503	0O2	C6-C5	-3.06	1.35	1.41
7	D	1503	0O2	O4'-C1'	-2.74	1.37	1.41
7	D	1503	0O2	C2'-C3'	-2.61	1.47	1.53
7	D	1503	0O2	O4'-C4'	-2.53	1.39	1.45

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	D	1503	0O2	C4'-O4'-C1'	-9.46	99.33	109.72
7	D	1503	0O2	PB-O3A-PA	-6.98	113.14	132.73
7	D	1503	0O2	PC-O3C-PD	-5.48	114.28	132.67
7	D	1503	0O2	N3-C2-N1	-3.07	122.77	127.44

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
7	D	1503	0O2	PB-O3B-PG	-2.43	124.52	132.67

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	D	1503	0O2	9	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	A	323/329 (98%)	-0.18	4 (1%) 81 73	0, 55, 172, 230	0
1	B	221/329 (67%)	-0.01	7 (3%) 51 39	0, 86, 193, 260	0
1	F	229/329 (69%)	0.02	13 (5%) 27 20	2, 123, 212, 293	0
1	G	217/329 (65%)	0.06	5 (2%) 64 53	5, 113, 204, 271	0
2	C	1335/1342 (99%)	-0.28	23 (1%) 73 63	0, 38, 168, 304	0
2	H	1335/1342 (99%)	-0.11	40 (2%) 54 42	0, 78, 206, 346	0
3	D	1160/1407 (82%)	-0.20	18 (1%) 74 65	0, 28, 152, 297	0
3	I	1160/1407 (82%)	-0.08	43 (3%) 45 35	0, 54, 183, 316	0
4	E	90/91 (98%)	-0.20	0 100 100	0, 33, 116, 167	0
4	J	76/91 (83%)	0.08	1 (1%) 79 71	12, 83, 181, 230	0
5	X	517/613 (84%)	-0.09	19 (3%) 45 35	0, 98, 238, 341	0
5	Y	458/613 (74%)	-0.07	18 (3%) 43 33	1, 100, 216, 296	0
All	All	7121/8222 (86%)	-0.13	191 (2%) 58 47	0, 63, 198, 346	0

The worst 5 of 191 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	I	10	ALA	9.8
2	H	982	GLY	9.5
2	H	981	ALA	8.6
3	I	521	LYS	8.5
2	H	983	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. 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, 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	LLDF	B-factors(\AA^2)	Q<0.9
6	ZN	I	1502	1/1	0.99	0.17	-0.72	49,49,49,49	0
6	ZN	D	1502	1/1	0.97	0.18	-0.94	8,8,8,8	0
6	ZN	D	1501	1/1	0.99	0.05	-1.65	54,54,54,54	0
6	ZN	I	1501	1/1	0.97	0.04	-1.83	60,60,60,60	0
7	0O2	D	1503	40/40	0.90	0.16	-1.91	20,20,20,20	0

6.5 Other polymers [i](#)

There are no such residues in this entry.