



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 10:39 PM GMT

PDB ID : 4YG2  
Title : X-ray crystal structur of Escherichia coli RNA polymerase sigma70 holoenzyme  
Authors : Murakami, K.S.  
Deposited on : 2015-02-25  
Resolution : 3.70 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  
<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

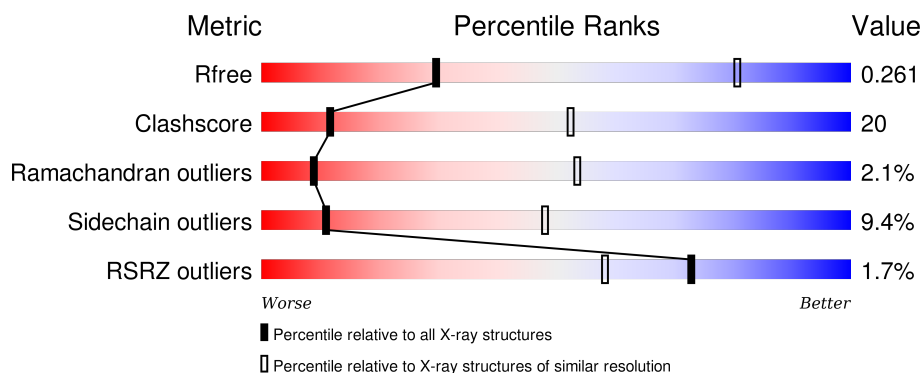
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.70 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	1101 (3.90-3.50)
Clashscore	102246	1224 (3.90-3.50)
Ramachandran outliers	100387	1172 (3.90-3.50)
Sidechain outliers	100360	1170 (3.90-3.50)
RSRZ outliers	91569	1108 (3.90-3.50)

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. 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>46%</div> <div>41%</div> <div>9%</div> <div>••</div> </div>
1	B	329	<div>2%</div> <div>37%</div> <div>27%</div> <div>•</div> <div>34%</div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	I	1342	<div><div></div><div>3%</div><div>59%</div><div>37%</div><div></div></div>
3	D	1407	<div><div></div><div>44%</div><div>33%</div><div>5%</div><div>17%</div><div></div></div>
3	J	1407	<div><div></div><div>%</div><div>45%</div><div>32%</div><div>5%</div><div>18%</div><div></div></div>
4	E	91	<div><div></div><div>62%</div><div>27%</div><div>9%</div><div></div></div>
4	K	91	<div><div></div><div>%</div><div>47%</div><div>33%</div><div>7%</div><div>13%</div><div></div></div>
5	F	613	<div><div></div><div>3%</div><div>41%</div><div>33%</div><div></div><div>24%</div><div></div></div>
5	L	613	<div><div></div><div>2%</div><div>40%</div><div>33%</div><div></div><div>23%</div><div></div></div>

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 55741 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	A	319	Total	C	N	O	S	0	0	0
			2490	1557	439	486	8			
1	B	217	Total	C	N	O	S	0	0	0
			1677	1047	295	329	6			
1	G	227	Total	C	N	O	S	0	0	0
			1755	1093	311	345	6			
1	H	216	Total	C	N	O	S	0	0	0
			1662	1038	292	326	6			

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1340	Total	C	N	O	S	0	0	0
			10570	6631	1841	2055	43			
2	I	1340	Total	C	N	O	S	0	0	0
			10566	6629	1840	2054	43			

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1166	Total	C	N	O	S	0	0	0
			9060	5697	1621	1696	46			
3	J	1155	Total	C	N	O	S	0	0	0
			9001	5659	1612	1684	46			

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	89	Total	C	N	O	S	0	0	0
			691	421	129	140	1			
4	K	79	Total	C	N	O	S	0	0	0
			627	382	118	126	1			

- Molecule 5 is a protein called RNA polymerase sigma factor RpoD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	468	Total	C	N	O	S	0	0	0
			3813	2389	678	723	23			
5	L	469	Total	C	N	O	S	0	0	0
			3821	2393	679	726	23			

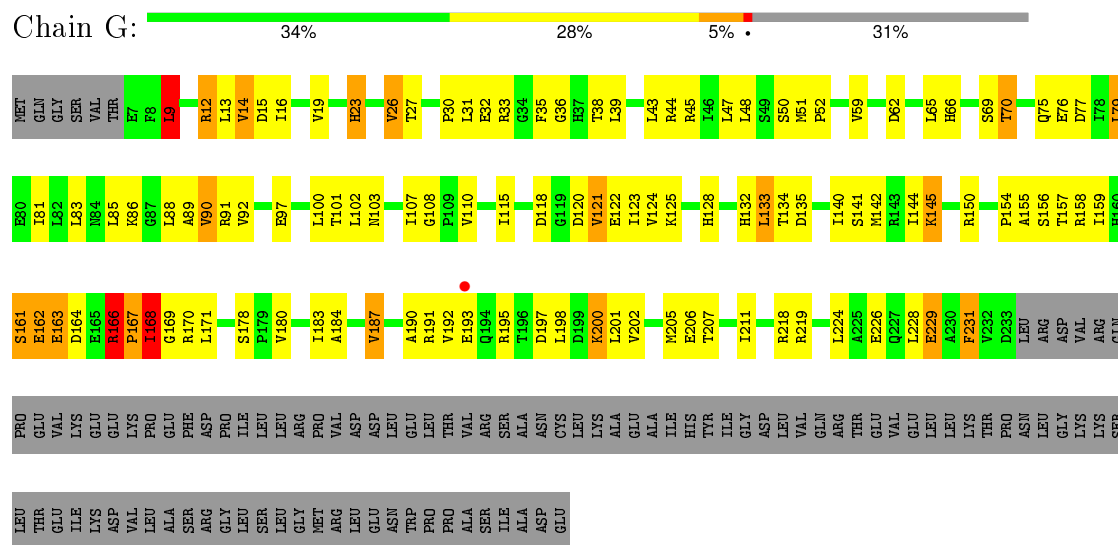
- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	J	1	Total	Mg	0	0
			1	1		
6	I	1	Total	Mg	0	0
			1	1		
6	D	1	Total	Mg	0	0
			1	1		
6	C	1	Total	Mg	0	0
			1	1		

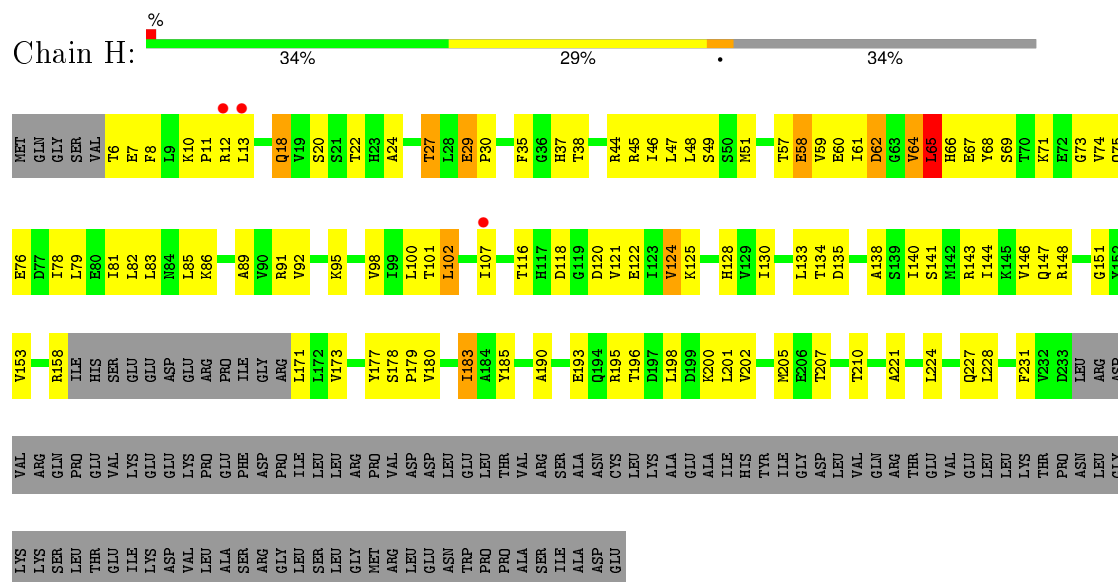
- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	J	2	Total	Zn	0	0
			2	2		
7	D	2	Total	Zn	0	0
			2	2		

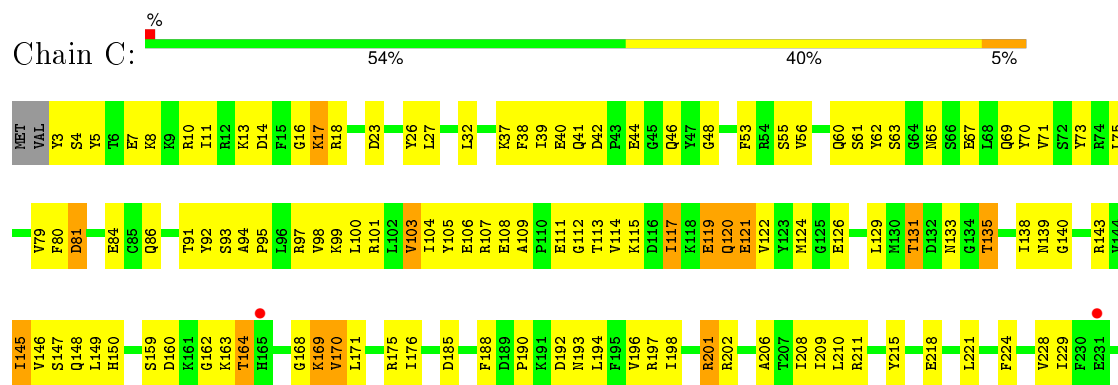


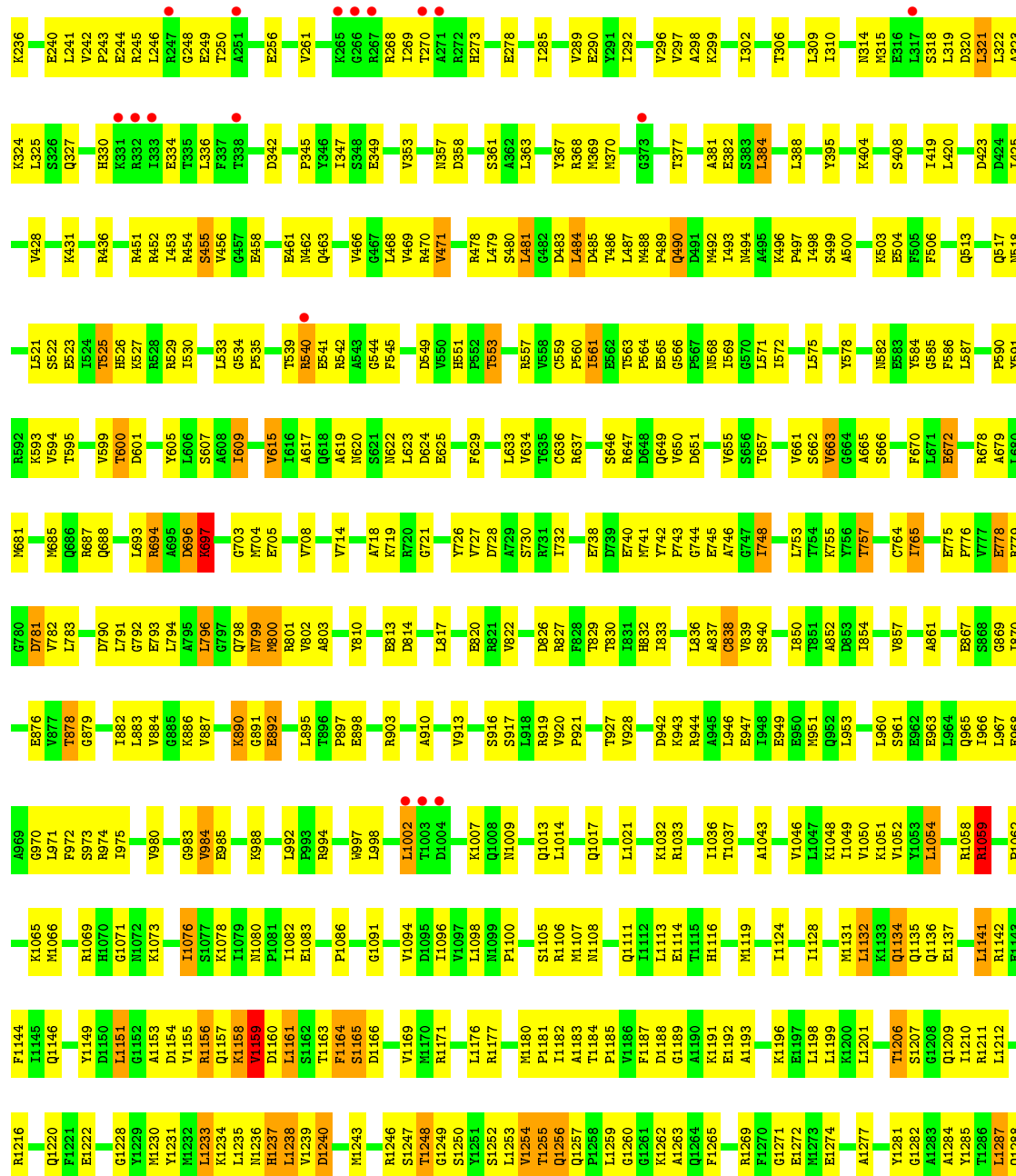


• Molecule 1: DNA-directed RNA polymerase subunit alpha

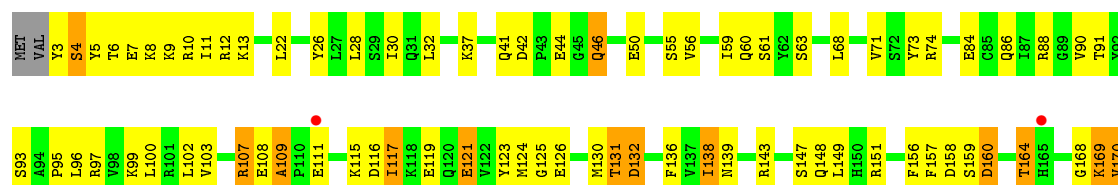


• Molecule 2: DNA-directed RNA polymerase subunit beta

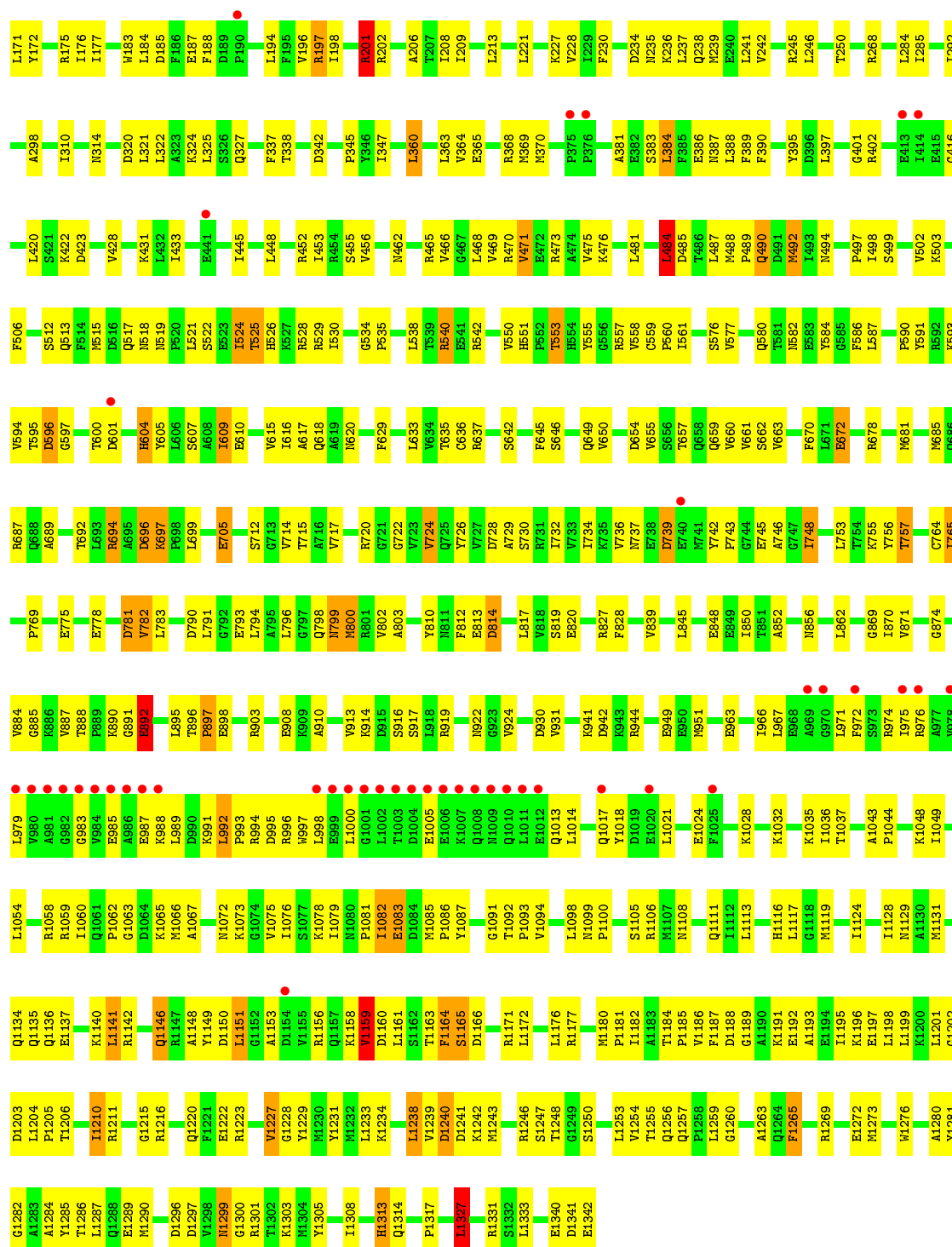




• Molecule 2: DNA-directed RNA polymerase subunit beta







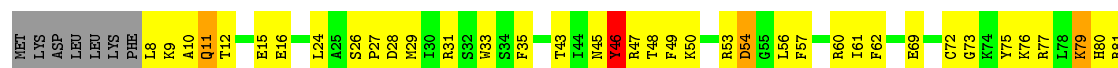
Chain D:

44%

33%

5%

17%

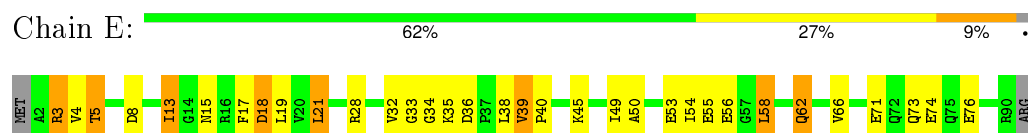


H1366	Q1279	R1203	THR	GLY	VAL	L796	E704	I611	V526	H430	V347	P247	Q164	G82
R1372	V1280	V1204	LYS	LYS	ALA	I797	T705	I612	V527	R431	D348	D247	Y165	V83
E1281	E1281	R1206	ASP	ASP	GLY	R798	T706	G613	L432	L433	Y349	L249	L169	C88
S1283	S1283	G1207	THR	THR	GLU	R799	T707	L614	Q433	Q434	S353	D248	E170	G89
A1374	A1374	G1208	PRO	PRO	THR	L800	T708	K615	A532	L435	S354	L252	E171	V90
V1285	V1285	V1209	ALA	ALA	VAL	V801	T709	P616	A533	Q436	I355	V253	F172	E91
K1286	K1286	I1210	LEU	LEU	VAL	V802	T710	A621	E534	F437	T356	P234	G173	V92
L1287	L1287	E1215	LYS	LYS	ASN	D806	G711	I624	R535	F438	V357	L255	D174	T93
A1288	A1288	I1211	ILE	ILE	TRP	L807	T712	I624	L536	P439	G358	R259	E175	Q94
N1289	N1289	V1140	VAL	VAL	ASP	L807	T713	I624	Y537	P439	P359	F176	F176	T95
A1294	A1294	I1220	ASP	ASP	PRO	T810	Q716	B634	L541	E443	V360	T262	K179	K96
K1297	K1297	I1221	ALA	ALA	HIS	E811	T717	S638	A542	G442	K361	S263	M180	V97
V1298	V1298	R1149	THR	THR	THR	D812	T718	V639	S543	K445	R362	R267	A184	A98
G1299	G1299	A1228	MET	MET	VAL	D813	T719	V639	L544	L446	L363	N266	E189	R99
A1300	A1300	V1229	VAL	VAL	VAL	I820	T720	I641	H545	Q447	Q365	D267	A189	R101
T1301	T1301	T1230	THR	THR	THR	N821	T721	I641	V548	L449	Q366	L268	A187	N102
ALA	ALA	R1231	SER	SER	SER	N822	T722	I641	K549	H450	G367	Y269	L188	G103
SER	SER	T1232	SER	SER	SER	T823	T723	I641	V550	H450	L368	R270	L189	L107
ALA	ALA	I1155	GLY	GLY	GLY	P824	T724	I641	R551	F461	A373	R278	K190	A108
ALA	ALA	I1156	LYS	LYS	LYS	V825	T725	I641	I552	G462	L374	L279	M192	A108
ALA	ALA	E1157	LYS	LYS	LYS	I826	T726	I641	E556	D464	E375	L291	E197	H113
ALA	ALA	A1158	THR	THR	THR	G828	T727	I641	D558	L472	L376	T291	E197	I114
ALA	ALA	S1160	THR	THR	THR	G829	T728	I641	A559	L473	F377	V292	E197	H115
ALA	ALA	G1161	THR	THR	THR	D830	T729	I641	N560	L474	P379	E295	L201	L120
ALA	ALA	I1162	THR	THR	THR	V831	T730	I641	G561	L475	E296	K296	E203	P121
ALA	ALA	V1163	THR	THR	THR	R832	T731	I641	G562	L476	Y352	R297	E207	S122
ALA	ALA	F1165	THR	THR	THR	R833	T732	I641	E562	L477	L385	N298	E207	R123
ALA	ALA	G1166	THR	THR	THR	N834	T733	I641	T567	L478	L386	L299	E207	R123
ALA	ALA	K1167	THR	THR	THR	L835	T734	I641	S568	E479	E396	L307	S210	L126
ALA	ALA	E1168	THR	THR	THR	D836	T735	I641	R570	N484	T392	L307	E211	L127
ALA	ALA	T1169	THR	THR	THR	R837	T736	I641	D571	N484	T393	L307	E211	L128
ALA	ALA	K1170	THR	THR	THR	V838	T737	I641	T572	N489	I394	R311	K213	P131
ALA	ALA	G1171	THR	THR	THR	R839	T738	I641	T573	L490	K395	R312	R214	L132
ALA	ALA	V1172	THR	THR	THR	V840	T739	I641	L579	L491	K396	R313	K215	L133
ALA	ALA	R1173	THR	THR	THR	R841	T740	I641	W580	N495	K398	R314	K216	R133
ALA	ALA	L1174	THR	THR	THR	V842	T741	I641	W581	N496	D410	T317	K217	D134
ALA	ALA	L1175	THR	THR	THR	R843	T742	I641	W582	N497	D411	T317	K218	L135
ALA	ALA	L1176	THR	THR	THR	V844	T743	I641	W583	N498	D412	T317	K219	E136
ALA	ALA	T1177	THR	THR	THR	R845	T744	I641	W584	N499	D413	T317	K220	R137
ALA	ALA	D1181	THR	THR	THR	V846	T745	I641	W585	N500	D414	T317	K221	V138
ALA	ALA	G1182	THR	THR	THR	R847	T746	I641	W586	N501	D415	T317	K222	L139
ALA	ALA	L1183	THR	THR	THR	V848	T747	I641	W587	N502	D416	T317	K223	Y140
ALA	ALA	E1184	THR	THR	THR	R849	T748	I641	W588	N503	D417	T317	K224	F141
ALA	ALA	P1185	THR	THR	THR	V850	T749	I641	W589	N504	D418	T317	K225	E142
ALA	ALA	Y1186	THR	THR	THR	R851	T750	I641	W590	N505	D419	T317	K226	S143
ALA	ALA	Y1187	THR	THR	THR	V852	T751	I641	W591	N506	D420	T317	K227	Y144
ALA	ALA	E1188	THR	THR	THR	R853	T752	I641	W592	N507	D421	T317	K228	L147
ALA	ALA	G1189	THR	THR	THR	V854	T753	I641	W593	N508	D422	T317	K229	E148
ALA	ALA	K1190	THR	THR	THR	R855	T754	I641	W594	N509	D423	T317	K230	R156
ALA	ALA	L1191	THR	THR	THR	V856	T755	I641	W595	N510	D424	T317	K231	Q157
ALA	ALA	Y1192	THR	THR	THR	R857	T756	I641	W596	N511	D425	T317	K232	T161
ALA	ALA	Y1193	THR	THR	THR	V858	T757	I641	W597	N512	D426	T317	K233	E162
ALA	ALA	R1194	THR	THR	THR	R859	T758	I641	W598	N513	D427	T317	K234	E163
ALA	ALA	Y1195	THR	THR	THR	V860	T759	I641	W599	N514	D428	T317	K235	Q157
ALA	ALA	Y1196	THR	THR	THR	R861	T760	I641	W600	N515	D429	T317	K236	T161
ALA	ALA	F1197	THR	THR	THR	V862	T761	I641	W601	N516	D430	T317	K237	E162
ALA	ALA	E1198	THR	THR	THR	R863	T762	I641	W602	N517	D431	T317	K238	E163
ALA	ALA	T1199	THR	THR	THR	V864	T763	I641	W603	N518	D432	T317	K239	E163
ALA	ALA	E1200	THR	THR	THR	R865	T764	I641	W604	N519	D433	T317	K240	E163
ALA	ALA	E1201	THR	THR	THR	V866	T765	I641	W605	N520	D434	T317	K241	E163
ALA	ALA	E1202	THR	THR	THR	R867	T766	I641	W606	N521	D435	T317	K242	E163
ALA	ALA	E1203	THR	THR	THR	V868	T767	I641	W607	N522	D436	T317	K243	E163
ALA	ALA	E1204	THR	THR	THR	R869	T768	I641	W608	N523	D437	T317	K244	E163
ALA	ALA	E1205	THR	THR	THR	V870	T769	I641	W609	N524	D438	T317	K245	E163
ALA	ALA	E1206	THR	THR	THR	R871	T770	I641	W610	N525	D439	T317	K246	E163
ALA	ALA	E1207	THR	THR	THR	V872	T771	I641	W611	N526	D440	T317	K247	E163
ALA	ALA	E1208	THR	THR	THR	R873	T772	I641	W612	N527	D441	T317	K248	E163
ALA	ALA	E1209	THR	THR	THR	V874	T773	I641	W613	N528	D442	T317	K249	E163
ALA	ALA	E1210	THR	THR	THR	R875	T774	I641	W614	N529	D443	T317	K250	E163
ALA	ALA	E1211	THR	THR	THR	V876	T775	I641	W615	N530	D444	T317	K251	E163
ALA	ALA	E1212	THR	THR	THR	R877	T776	I641	W616	N531	D445	T317	K252	E163
ALA	ALA	E1213	THR	THR	THR	V878	T777	I641	W617	N532	D446	T317	K253	E163
ALA	ALA	E1214	THR	THR	THR	R879	T778	I641	W618	N533	D447	T317	K254	E163
ALA	ALA	E1215	THR	THR	THR	V880	T779	I641	W619	N534	D448	T317	K255	E163
ALA	ALA	E1216	THR	THR	THR	R881	T780	I641	W620	N535	D449	T317	K256	E163
ALA	ALA	E1217	THR	THR	THR	V882	T781	I641	W621	N536	D450	T317	K257	E163
ALA	ALA	E1218	THR	THR	THR	R883	T782	I641	W622	N537	D451	T317	K258	E163
ALA	ALA	E1219	THR	THR	THR	V884	T783	I641	W623	N538	D452	T317	K259	E163
ALA	ALA	E1220	THR	THR	THR	R885	T784	I641	W624	N539	D453	T317	K260	E163
ALA	ALA	E1221	THR	THR	THR	V886	T785	I641	W625	N540	D454	T317	K261	E163
ALA	ALA	E1222	THR	THR	THR	R887	T786	I641	W626	N541	D455	T317	K262	E163
ALA	ALA	E1223	THR	THR	THR	V888	T787	I641	W627	N542	D456	T317	K263	E163
ALA	ALA	E1224	THR	THR	THR	R889	T788	I641	W628	N543	D457	T317	K264	E163
ALA	ALA	E1225	THR	THR	THR	V890	T789	I641	W629	N544	D458	T317	K265	E163
ALA	ALA	E1226	THR	THR	THR	R891	T790	I641	W630	N545	D459	T317	K266	E163
ALA	ALA	E1227	THR	THR	THR	V892	T791	I641	W631	N546	D460	T317	K267	E163
ALA	ALA	E1228	THR	THR	THR	R893	T792	I641	W632	N547	D461	T317	K268	E163
ALA	ALA	E1229	THR	THR	THR	V894	T793	I641	W633	N548	D462	T317	K269	E163
ALA	ALA	E1230	THR	THR	THR	R895	T794	I641	W634	N549	D463	T317	K270	E163
ALA	ALA	E1231	THR	THR	THR	V896	T795	I641	W635	N550	D464	T317	K271	E163
ALA	ALA	E1232	THR	THR	THR	R897	T796	I641	W636	N551	D465	T317	K272	E163
ALA	ALA	E1233	THR	THR	THR	V898	T797	I641	W637	N552	D466	T317	K273	E163
ALA	ALA	E1234	THR	THR	THR	R899	T798	I641	W638	N553	D467	T317	K274	E163
ALA	ALA	E1235	THR	THR	THR	V900	T799	I641	W639	N554	D468	T317	K275	E163
ALA	ALA	E1236	THR	THR	THR	R901	T800	I641	W640	N555	D469	T317	K276	E163
ALA	ALA	E1237	THR	THR	THR	V902	T801	I641	W641	N556	D470	T317	K277	E163
ALA	ALA	E1238	THR	THR	THR	R903	T802	I641	W642	N557	D471	T317	K278	E163
ALA	ALA	E1239	THR	THR	THR	V904	T803	I641	W643	N558	D472	T317	K279	E163
ALA	ALA	E1240	THR	THR	THR	R905	T804	I641	W644	N559	D473	T317	K280	E163
ALA	ALA	E1241	THR	THR	THR	V906	T805	I641	W645	N560	D474	T317	K281	E163
ALA	ALA	E1242	THR	THR	THR	R907	T806	I641	W646	N561	D475	T317	K282	E163
ALA	ALA	E1243	THR	THR	THR	V908	T807	I641	W647	N562	D476	T317	K283	E163
ALA	ALA	E1244	THR	THR	THR	R909	T808	I641	W648	N563	D477	T317	K284	E163

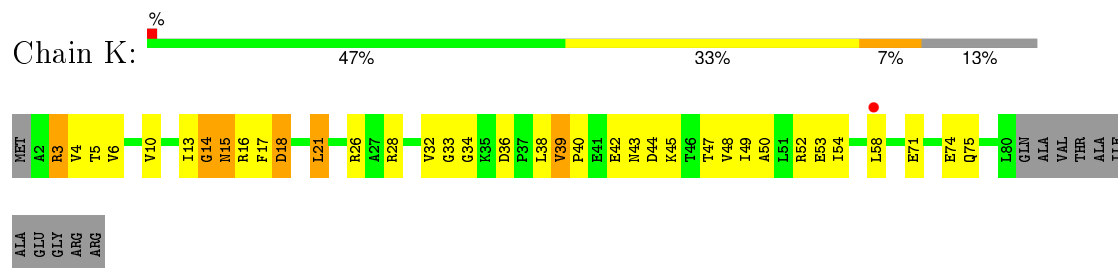
Chain J:  45% 32% 5% 18%



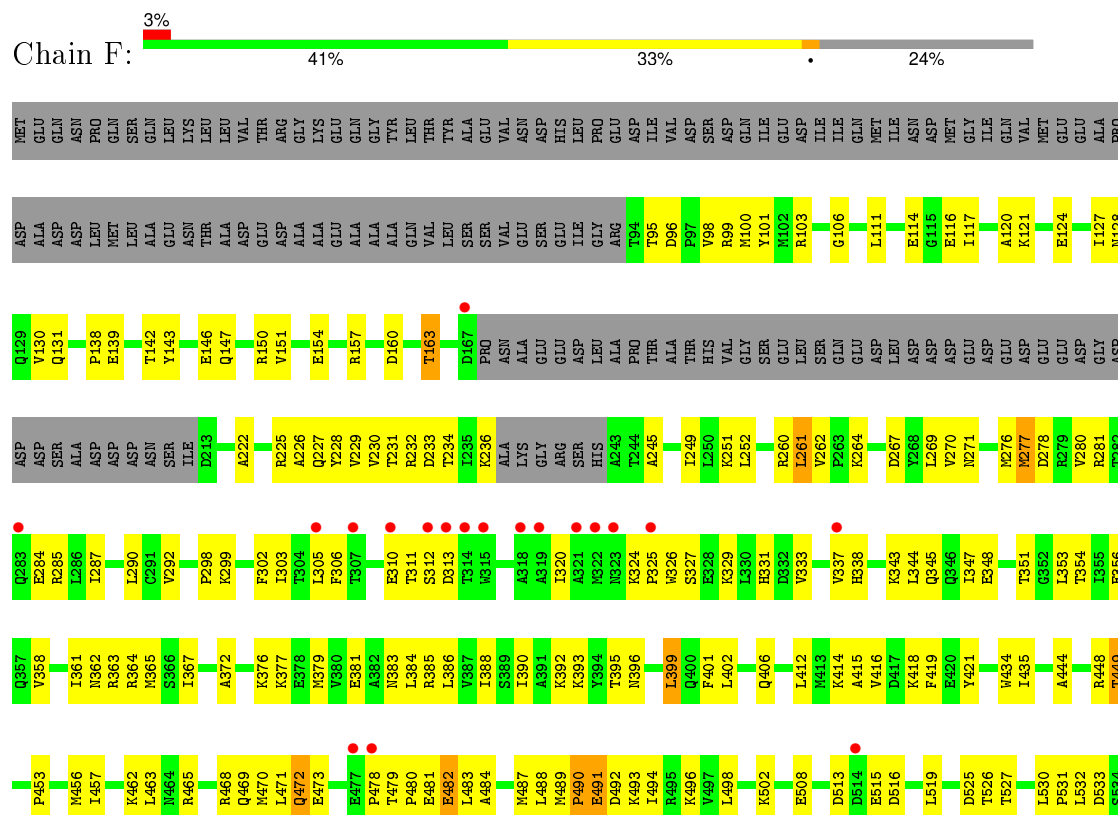
- Molecule 4: DNA-directed RNA polymerase subunit omega



- Molecule 4: DNA-directed RNA polymerase subunit omega



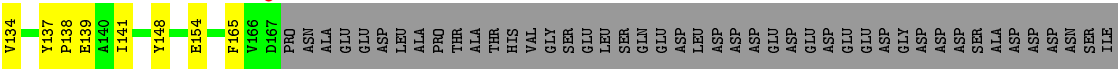
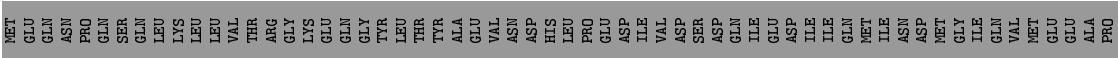
- Molecule 5: RNA polymerase sigma factor RpoD





ASP

• Molecule 5: RNA polymerase sigma factor RpoD



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	187.31Å 205.90Å 309.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.98 – 3.70 29.98 – 3.70	Depositor EDS
% Data completeness (in resolution range)	96.8 (29.98-3.70) 92.5 (29.98-3.70)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.49 (at 3.75Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.191 , 0.260 0.194 , 0.261	Depositor DCC
$R_{free}$ test set	5897 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	114.6	Xtriage
Anisotropy	0.143	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 61.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.25$	Xtriage
Outliers	2 of 123448 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	55741	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	140.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.66% 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, 0.375 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, 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	A	0.62	0/2524	0.81	2/3421 (0.1%)
1	B	0.58	0/1697	0.84	0/2300
1	G	0.57	0/1777	0.78	0/2408
1	H	0.51	0/1681	0.83	3/2278 (0.1%)
2	C	0.69	2/10739 (0.0%)	0.86	7/14489 (0.0%)
2	I	0.57	0/10735	0.77	3/14484 (0.0%)
3	D	0.70	0/9200	0.90	12/12423 (0.1%)
3	J	0.64	0/9140	0.83	5/12341 (0.0%)
4	E	0.66	0/693	0.82	0/935
4	K	0.61	0/629	0.77	0/847
5	F	0.56	0/3864	0.81	3/5194 (0.1%)
5	L	0.52	0/3872	0.71	1/5205 (0.0%)
All	All	0.63	2/56551 (0.0%)	0.82	36/76325 (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	C	0	2
2	I	0	2
3	D	0	2
3	J	0	3
5	F	0	2
All	All	0	11

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	949	GLU	CB-CG	6.52	1.64	1.52

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	838	CYS	CB-SG	-6.02	1.72	1.82

The worst 5 of 36 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	I	484	LEU	CA-CB-CG	9.03	136.06	115.30
2	C	796	LEU	CB-CG-CD2	-8.02	97.36	111.00
2	C	1161	LEU	CA-CB-CG	-7.43	98.22	115.30
3	D	268	LEU	CA-CB-CG	-7.05	99.09	115.30
2	C	544	GLY	N-CA-C	-6.83	96.01	113.10

There are no chirality outliers.

5 of 11 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	109	ALA	Peptide
2	C	236	LYS	Peptide
3	D	1184	ASP	Peptide
3	D	901	ARG	Peptide
5	F	600	HIS	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	A	2490	0	2542	129	0
1	B	1677	0	1703	75	0
1	G	1755	0	1773	96	0
1	H	1662	0	1687	79	0
2	C	10570	0	10582	460	0
2	I	10566	0	10576	403	1
3	D	9060	0	9209	451	1
3	J	9001	0	9168	412	0
4	E	691	0	695	24	0
4	K	627	0	634	27	0
5	F	3813	0	3880	150	0
5	L	3821	0	3884	163	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	C	1	0	0	0	0
6	D	1	0	0	0	0
6	I	1	0	0	0	0
6	J	1	0	0	0	0
7	D	2	0	0	0	0
7	J	2	0	0	0	0
All	All	55741	0	56333	2271	1

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 20.

The worst 5 of 2271 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)
1:G:190:ALA:HB2	1:G:200:LYS:HB2	1.42	1.01
1:A:296:GLY:H	1:A:299:SER:HB2	1.26	1.00
5:L:484:ALA:HB1	5:L:491:GLU:HB2	1.41	0.99
3:D:210:SER:HB2	3:D:213:LYS:HB2	1.47	0.97
3:D:1183:SER:HA	3:J:206:ASN:HD21	1.29	0.96

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:D:712:GLN:NE2	2:I:862:LEU:O[4_545]	2.14	0.06

## 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	317/329 (96%)	246 (78%)	53 (17%)	18 (6%)	<b>2</b> 28

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	213/329 (65%)	193 (91%)	16 (8%)	4 (2%)	10	55
1	G	225/329 (68%)	198 (88%)	18 (8%)	9 (4%)	4	38
1	H	212/329 (64%)	196 (92%)	12 (6%)	4 (2%)	10	55
2	C	1338/1342 (100%)	1205 (90%)	112 (8%)	21 (2%)	12	58
2	I	1338/1342 (100%)	1197 (90%)	120 (9%)	21 (2%)	12	58
3	D	1162/1407 (83%)	1030 (89%)	105 (9%)	27 (2%)	8	52
3	J	1151/1407 (82%)	1027 (89%)	98 (8%)	26 (2%)	8	52
4	E	87/91 (96%)	81 (93%)	5 (6%)	1 (1%)	17	65
4	K	77/91 (85%)	71 (92%)	3 (4%)	3 (4%)	4	38
5	F	462/613 (75%)	423 (92%)	32 (7%)	7 (2%)	13	59
5	L	463/613 (76%)	425 (92%)	30 (6%)	8 (2%)	11	57
All	All	7045/8222 (86%)	6292 (89%)	604 (9%)	149 (2%)	9	53

5 of 149 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	14	VAL
1	A	107	ILE
1	A	136	GLU
1	A	162	GLU
1	A	167	PRO

### 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	278/286 (97%)	227 (82%)	51 (18%)	2	14
1	B	186/286 (65%)	171 (92%)	15 (8%)	15	54
1	G	193/286 (68%)	164 (85%)	29 (15%)	3	26
1	H	183/286 (64%)	170 (93%)	13 (7%)	18	60
2	C	1155/1157 (100%)	1053 (91%)	102 (9%)	12	51

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	I	1154/1157 (100%)	1053 (91%)	101 (9%)	12	51
3	D	962/1168 (82%)	876 (91%)	86 (9%)	12	51
3	J	960/1168 (82%)	873 (91%)	87 (9%)	12	49
4	E	72/75 (96%)	63 (88%)	9 (12%)	6	33
4	K	67/75 (89%)	60 (90%)	7 (10%)	9	42
5	F	417/540 (77%)	385 (92%)	32 (8%)	16	56
5	L	418/540 (77%)	379 (91%)	39 (9%)	11	49
All	All	6045/7024 (86%)	5474 (91%)	571 (9%)	11	48

5 of 571 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	E	21	LEU
1	G	200	LYS
4	K	18	ASP
5	F	154	GLU
5	F	572	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 64 such sidechains are listed below:

Mol	Chain	Res	Type
5	F	469	GLN
2	I	620	ASN
5	L	258	GLN
5	F	545	HIS
1	G	66	HIS

### 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 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

## 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 ⓘ

### 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	319/329 (96%)	-0.33	1 (0%) 94 90	84, 129, 182, 196	0
1	B	217/329 (65%)	-0.16	5 (2%) 64 48	91, 164, 203, 214	0
1	G	227/329 (68%)	-0.25	1 (0%) 93 88	126, 154, 181, 218	0
1	H	216/329 (65%)	-0.07	3 (1%) 78 64	118, 174, 198, 207	0
2	C	1340/1342 (99%)	-0.35	19 (1%) 78 64	58, 114, 218, 253	0
2	I	1340/1342 (99%)	-0.16	45 (3%) 49 34	93, 148, 222, 378	0
3	D	1166/1407 (82%)	-0.36	1 (0%) 95 94	63, 107, 168, 218	0
3	J	1155/1407 (82%)	-0.24	14 (1%) 81 68	85, 128, 189, 228	0
4	E	89/91 (97%)	-0.56	0 100 100	67, 116, 140, 152	0
4	K	79/91 (86%)	-0.24	1 (1%) 79 66	99, 139, 205, 222	0
5	F	468/613 (76%)	-0.14	19 (4%) 41 27	93, 160, 273, 305	0
5	L	469/613 (76%)	-0.13	13 (2%) 56 41	114, 172, 279, 307	0
All	All	7085/8222 (86%)	-0.25	122 (1%) 73 58	58, 135, 209, 378	0

The worst 5 of 122 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	I	982	GLY	11.0
5	F	319	ALA	6.7
2	I	983	GLY	5.3
5	L	318	ALA	5.0
2	I	1004	ASP	5.0

### 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 [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, 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
7	ZN	D	2003	1/1	0.99	0.17	-0.64	94,94,94,94	0
7	ZN	D	2002	1/1	0.99	0.09	-1.07	113,113,113,113	0
7	ZN	J	2003	1/1	0.98	0.11	-1.08	94,94,94,94	0
7	ZN	J	2002	1/1	0.97	0.05	-1.43	105,105,105,105	0
6	MG	J	2001	1/1	0.96	0.25	-	94,94,94,94	0
6	MG	C	1401	1/1	0.95	0.51	-	94,94,94,94	0
6	MG	I	1401	1/1	0.96	0.44	-	94,94,94,94	0
6	MG	D	2001	1/1	0.96	0.35	-	94,94,94,94	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.