



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:25 AM GMT

PDB ID : 2A6E
Title : Crystal structure of the T. Thermophilus RNA polymerase holoenzyme
Authors : Artsimovitch, I.; Vassilyeva, M.N.; Svetlov, D.; Svetlov, V.; Perederina, A.;
Igarashi, N.; Matsugaki, N.; Wakatsuki, S.; Tahirov, T.H.; Vassilyev, D.G.;
RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on : 2005-07-02
Resolution : 2.80 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

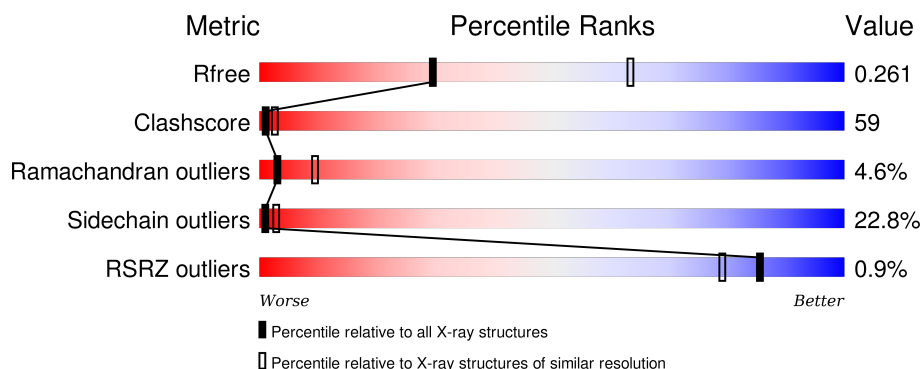
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	315	
1	B	315	
1	K	315	
1	L	315	
2	C	1119	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	M	1119	<div><div><div></div><div></div><div></div></div><div><div>%</div><div>25%</div><div>57%</div><div>17%</div><div></div></div></div>
3	D	1524	<div><div><div></div><div></div><div></div></div><div><div>%</div><div>23%</div><div>51%</div><div>17%</div><div>9%</div><div></div></div></div>
3	N	1524	<div><div><div></div><div></div><div></div></div><div><div>%</div><div>24%</div><div>51%</div><div>15%</div><div>9%</div><div></div></div></div>
4	E	99	<div><div><div></div><div></div><div></div></div><div><div></div><div>22%</div><div>52%</div><div>20%</div><div></div></div></div>
4	O	99	<div><div><div></div><div></div><div></div></div><div><div>%</div><div>28%</div><div>48%</div><div>19%</div><div></div></div></div>
5	F	423	<div><div><div></div><div></div><div></div></div><div><div>%</div><div>19%</div><div>47%</div><div>14%</div><div>18%</div><div></div></div></div>
5	P	423	<div><div><div></div><div></div><div></div></div><div><div>%</div><div>22%</div><div>49%</div><div>10%</div><div>18%</div><div></div></div></div>

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 58679 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 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
1	B	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
1	K	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
1	L	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1119	Total	C	N	O	S	0	0	0
			8829	5581	1577	1647	24			
2	M	1119	Total	C	N	O	S	0	0	0
			8829	5581	1577	1647	24			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1392	Total	C	N	O	S	0	0	0
			10797	6819	1925	2020	33			
3	N	1392	Total	C	N	O	S	0	0	0
			10797	6819	1925	2020	33			

- Molecule 4 is a protein called RNA polymerase omega chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	95	Total	C	N	O	S	0	0	0
			769	488	133	144	4			
4	O	95	Total	C	N	O	S	0	0	0
			769	488	133	144	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	345	Total	C	N	O	S	0	0	0
			2771	1744	504	519	4			
5	P	345	Total	C	N	O	S	0	0	0
			2771	1744	504	519	4			

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

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

- Molecule 7 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	D	1	Total	Mg	0	0
			1	1		
7	N	1	Total	Mg	0	0
			1	1		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	191	Total	O	0	0
			191	191		
8	B	181	Total	O	0	0
			181	181		
8	C	767	Total	O	0	0
			767	767		
8	D	1100	Total	O	0	0
			1100	1100		
8	E	93	Total	O	0	0
			93	93		
8	F	333	Total	O	0	0
			333	333		
8	K	151	Total	O	0	0
			151	151		
8	L	179	Total	O	0	0
			179	179		
8	M	739	Total	O	0	0
			739	739		

Continued on next page...

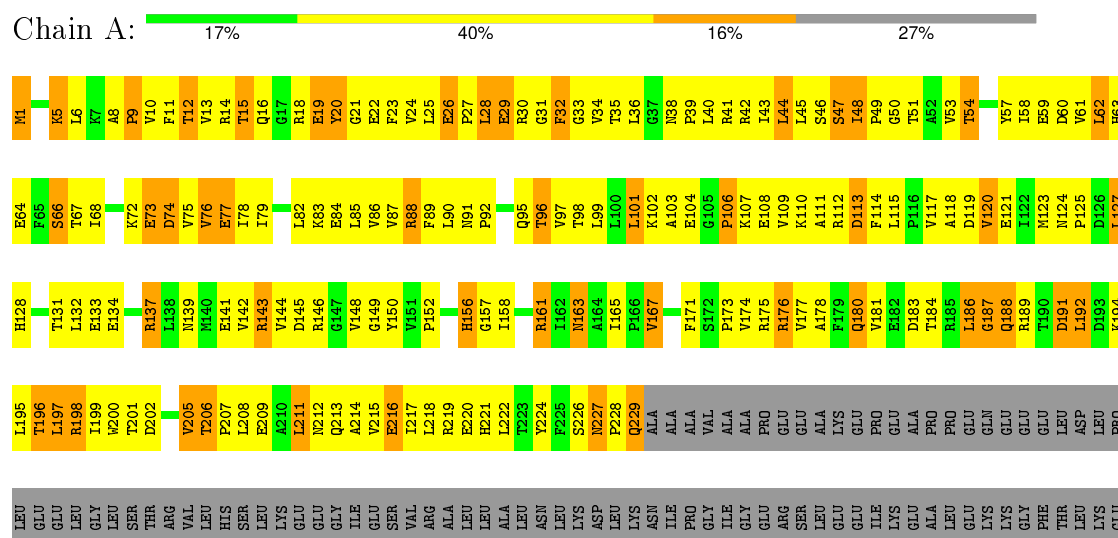
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	N	1038	Total 1038	O 1038	0	0
8	O	78	Total 78	O 78	0	0
8	P	267	Total 267	O 267	0	0

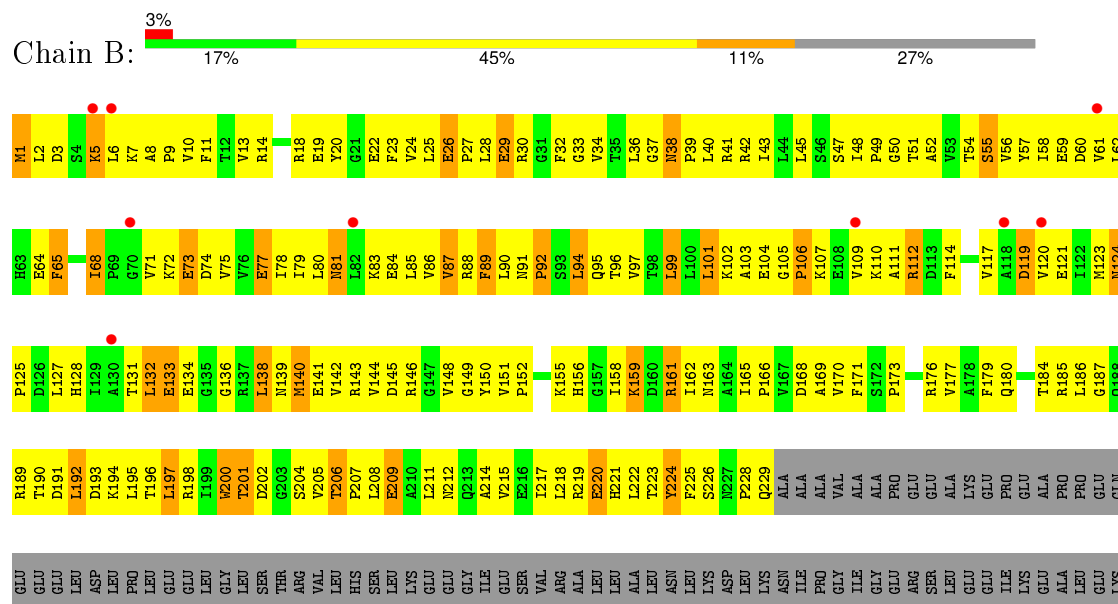
3 Residue-property plots

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.

• Molecule 1: DNA-directed RNA polymerase alpha chain

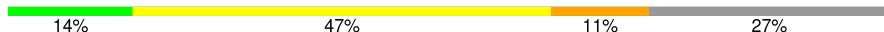


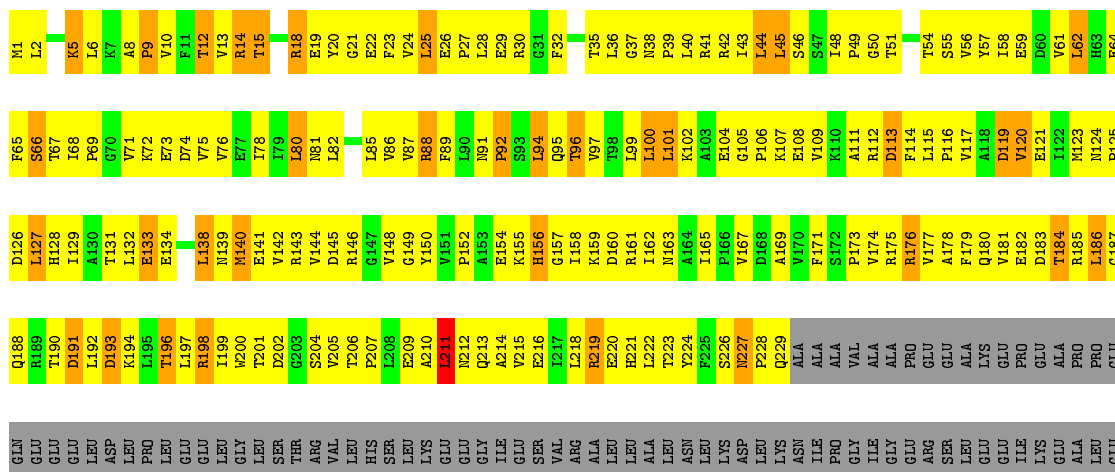
• Molecule 1: DNA-directed RNA polymerase alpha chain



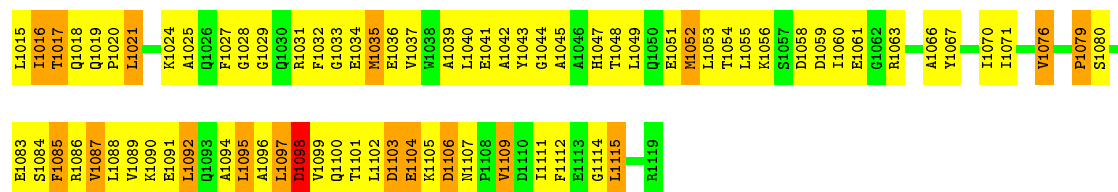
LYS
GLY
PHE
THR
LEU
LEU
LYS
GLU

• Molecule 1: DNA-directed RNA polymerase alpha chain

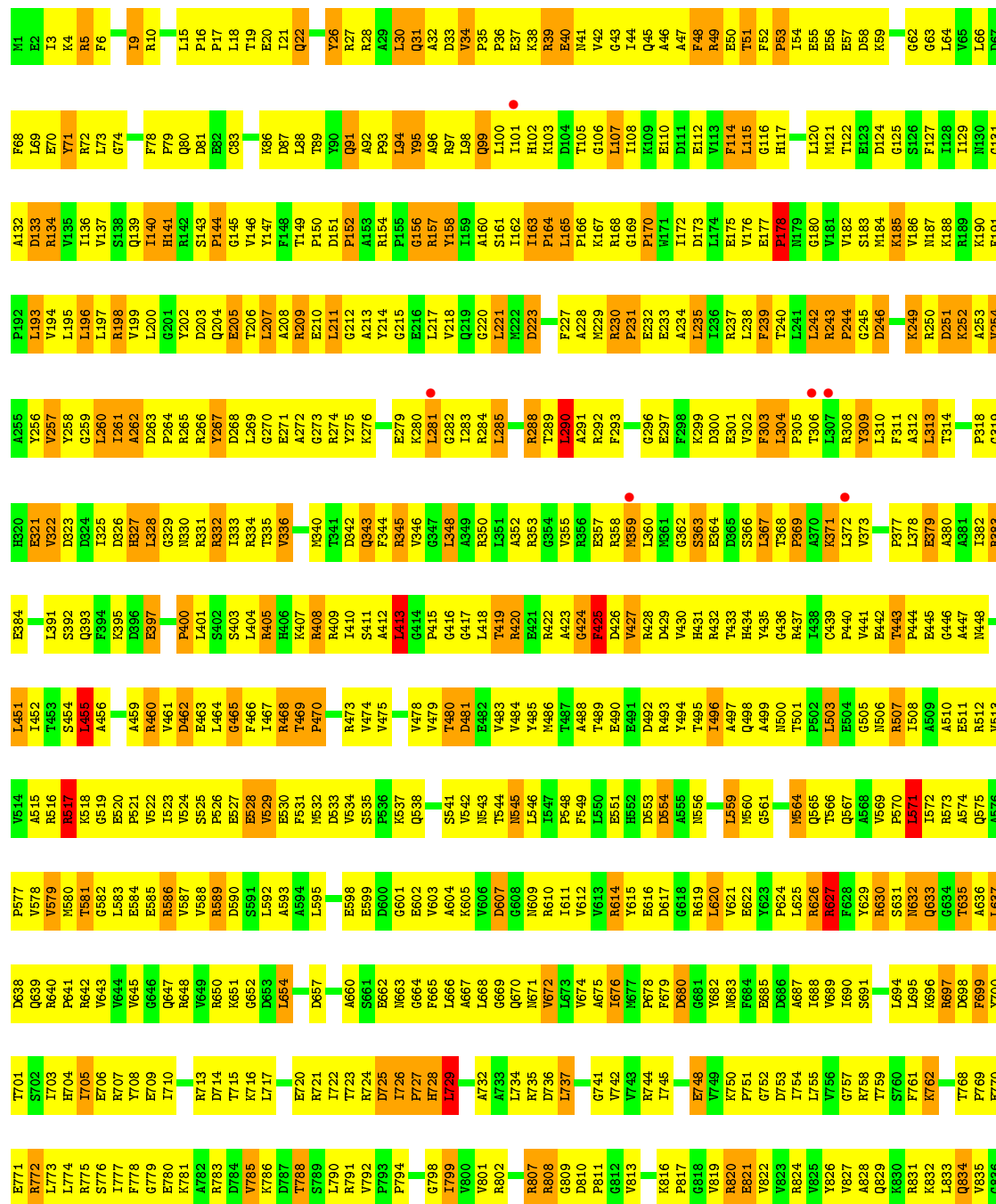
Chain K: 

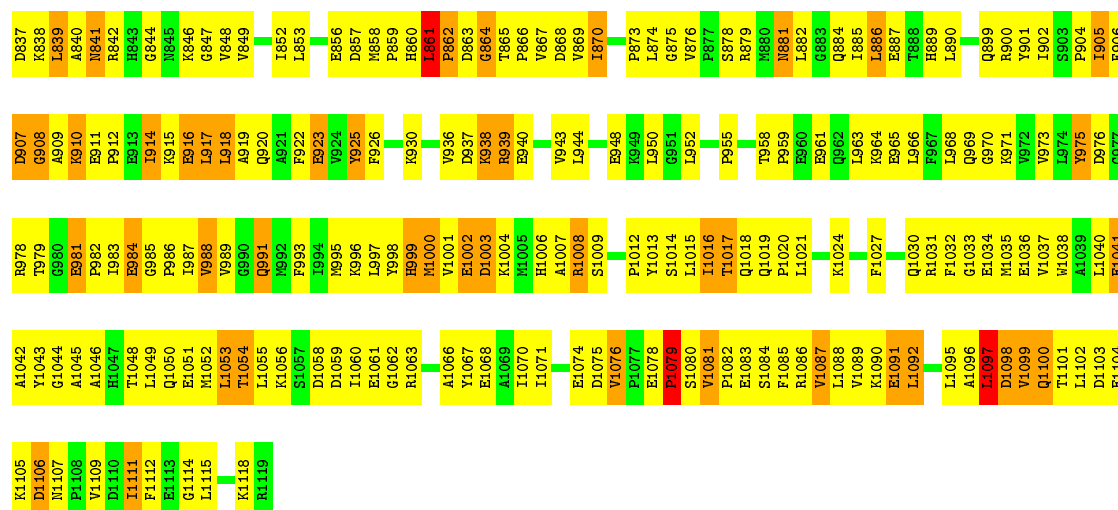




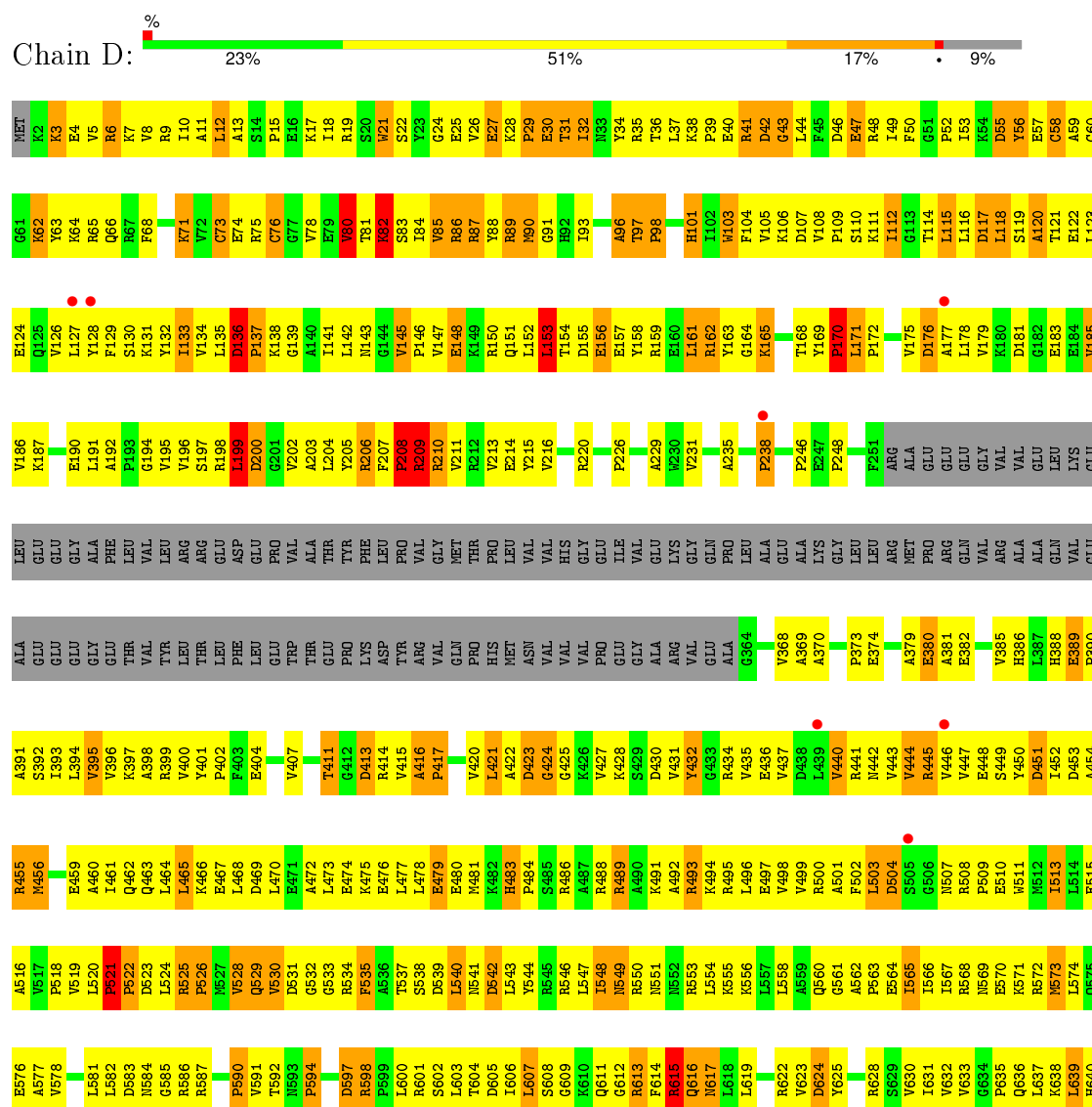


• Molecule 2: DNA-directed RNA polymerase beta chain





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



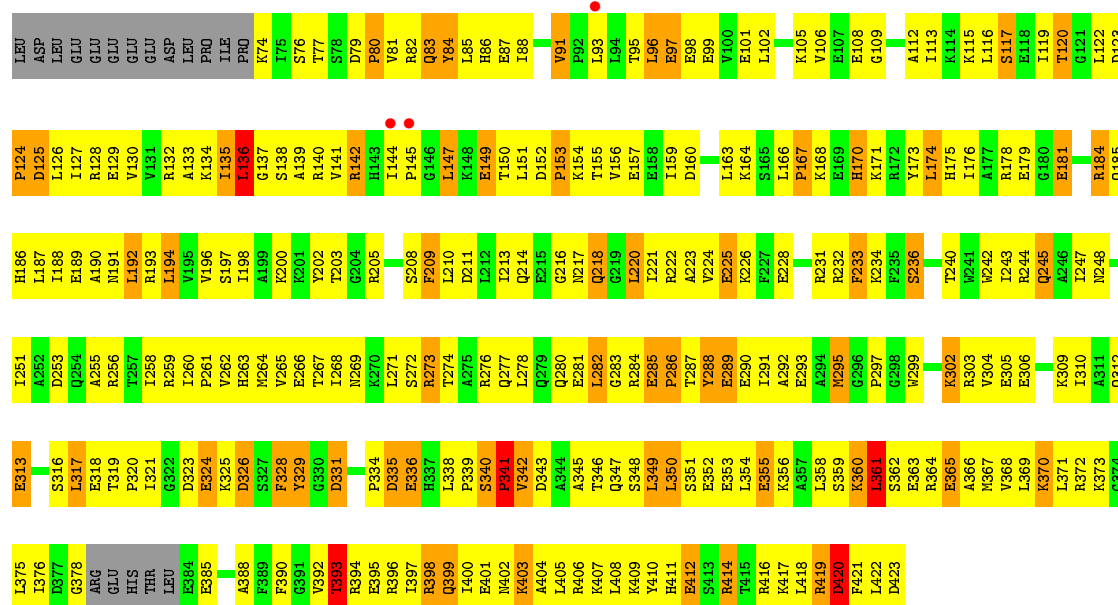
Q1477	G1478	T1413	E1351	A1220	L1156	L1094	Q1093	L964	Y900	Q831	N768	N703	Q641
S1478	D1479	P1414	I1352	I1221	G1157	L1095	Q1094	E965	Q901	R832	L769	R704	C642
F1480	E1481	V1415	Q1353	G1222	Y1158	R1096	I1035	E966	I902	E933	L770	R705	G643
V1481	G1482	A1416	D1291	I1223	R1159	K1097	I1036	A967	D903	T834	S771	P706	L644
R1482	K1417	V1292	V1292	V1224	L1098	L1098	Q1037	D968	V904	S835	P772	T707	P645
F1483	K1418	E1293	E1293	E1224	E1161	V1099	L1038	R969	P905	V836	A773	L708	K646
T1484	Q1359	V1294	E1295	G1230	E1162	D1100	C1039	K970	G906	G837	S774	H709	R647
Q1485	Q1360	E1296	E1296	E1231	G1163	V1101	G1040	L971	E907	R838	R710	R710	H648
V1486	G1360	P1232	P1232	E1232	R1164	L1102	L1041	L972	R908	L839	E776	L711	A649
F1487	F1361	G1233	G1233	G1233	Y1165	H1103	G1042	Q973	N909	R840	P777	G712	L650
D1488	K1301	Q1235	Q1235	Q1235	L1166	E1104	G1043	Q974	S910	Y841	L778	I713	E651
T1491	E1302	L1236	L1236	L1236	S1167	L1105	M1044	E975	I911	V842	A779	Q714	L652
S1427	Y1303	T1237	T1237	T1237	D1169	V1107	Q1046	Q976	K912	F843	K780	W715	F653
A1428	K1304	R1238	R1238	R1238	D1170	E1108	K1047	Y978	I914	A844		F716	K654
L1429	L1305	T1240	T1240	T1240	L1171	E1109	P1048	E979	N915	D847	P718	Q717	P655
S1430	P1306	F1241	F1241	F1241	D1111	G1050	S1049	N980	Y916	E848	L785	W719	L657
T1431	K1307	H1242	H1242	H1242	G1112	E1051	E1051	D885	Q917	A849	I786	L720	L658
K1432	R1310	T1243	T1243	T1243	G1113	T1052	T1052	R986	N924	R857	R783	Q727	L657
S1433	L1311	G1244	G1244	G1244	T1114	F1053	F1053	E987	Y920	E852	L789	W721	L659
W1434	V1313	G1245	G1245	G1245	N1116	V1055	V1055	R988	R921	V853	Y790	E722	K660
A1438	K1314	G1248	G1248	G1248	E1182	P1056	P1056	I992	G923	I857	I792	W726	E663
S1439	D1315	A1249	A1249	A1249	L1183	V1057	V1057	I993	N925	V858	T793	Q727	L666
F1440	G1316	D1250	D1250	D1250	Q1184	R1058	R1058	Q994	E925	D859	Q794	L728	A667
Q1441	D1317	I1251	I1251	I1251	E1185	S1059	S1059	Q995	R926	L860	W795	H729	P668
N1442	V1318	T1252	T1252	T1252	L1186	P1120	P1120	N996	T927	K861	R796	P730	N669
T1443	E1380	L1253	L1253	L1253	P1187	F1061	F1061	R997	A928	D862	K797	W731	V670
L1444	V1381	Q1254	Q1254	Q1254	L1188	F1123	F1123	E998	R929	V863	E798		K671
H1445	T1382	G1255	G1255	G1255	R1189	Q1124	Q1124	T999	L930	R864	K799	E734	A672
V1446	D1383	L1256	L1256	L1256	S1190	P1125	P1125	T1000	I931	T865	K800	F735	A673
L1447	Q1322	T1257	T1257	T1257	L1191	L1065	L1065	E1001	Y936	R866		F736	K674
T1448	G1323	R1258	R1258	R1258	L1192	T1066	T1066	Y1002	Y937	R867	L804	A738	K676
I1452	L1325	V1259	V1259	V1259	T1129	L1068	L1068	V1003	G938	M869		D739	L677
A1453	K1326	I1260	I1260	I1260	Q1195	E1069	E1069	A1006	F939	F806	F806	F740	E678
G1454	R1327	E1261	E1261	E1261	S1131	Y1070	Y1070	V1007	T940	E874	A807	D741	R679
K1455	G1328	F1263	F1263	F1263	L1132	F1071	F1071	F1008	S941	T875	T808		Q680
D1456	I1330	E1264	E1264	E1264	R1133	I1072	I1072	K1009	S942	S876	P809	Q744	R681
E1457	D1331	A1265	A1265	A1265	V1200	R1134	R1134	N1010	T943	P877	E810	M745	D682
E1458	P1332	R1266	R1266	R1266	C1201	K1136	K1136	F1011	T944	G878	E811		L683
L1459	H1333	Q1267	Q1267	Q1267	Q1202	R1137	R1137	E1012	S945	R879	A812	V749	K684
I1460	Q1334	P1268	P1268	P1268	A1138	G1076	G1076	E1013	G946	I800	L813		D685
G1461	L1335	K1269	K1269	K1269	D1139	R1077	R1077	H1014	Y947		L814	S752	E686
L1462	E1336	A1270	A1270	A1270	Y1205	K1078	K1078	Y1015	T948	A883		S753	V687
K1463	L1337	K1271	K1271	K1271	G1206			P1019	T949	R884	E817	S754	K688
E1464	E1338	A1272	A1272	A1272	E1141	A1142	A1142	L1020	G950	E888	R818	A755	D689
N1465	K1339	V1273	V1273	V1273	D1208	G1143	G1143	Y1021	I951	E889	G819	Q756	A690
V1466	G1340	I1274	I1274	I1274	L1209	L1144	L1144	D952	D952	A889	P820	A757	L691
I1467	P1341	T1277	T1277	T1277	S1210	Y1145	Y1145	Y955	V955	V890	V821	E758	E692
L1468	E1342	D1278	D1278	D1278	M1211	G1146	G1146	M1022	I956	E891	A822	A759	E693
G1469	A1343	G1279	G1279	G1279	A1212	R1147	R1147	A1024	P957	D892	L823	R760	V694
K1470	R1344	V1280	V1280	V1280	R1213	V1148	V1148	Q1025	P958	E893	K824	I761	L695
L1471	E1345	V1281	V1281	V1281	P1214	T1088	T1088	S1026	E958	K825	A825	Q762	H696
I1472	K1346	V1281	V1281	V1281	V1215	A1089	A1089	G1027	E959	V895	P826	W763	G697
A1499	Y1347	R1282	R1282	R1282	G1216	D1090	D1090	A1028	K960	A896	L827	L764	K698
E1410	L1348	I1217	I1217	I1217	E1152	V1153	V1153	R1029	R961	R897	R828	S765	K699
L1411	V1349	E1284	E1284	E1284	E1154	G1092	G1092		Y963	E898	A830		
K1412	E1350	E1285	E1285	E1285	L1155	Y1093	Y1093			L899			L702

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

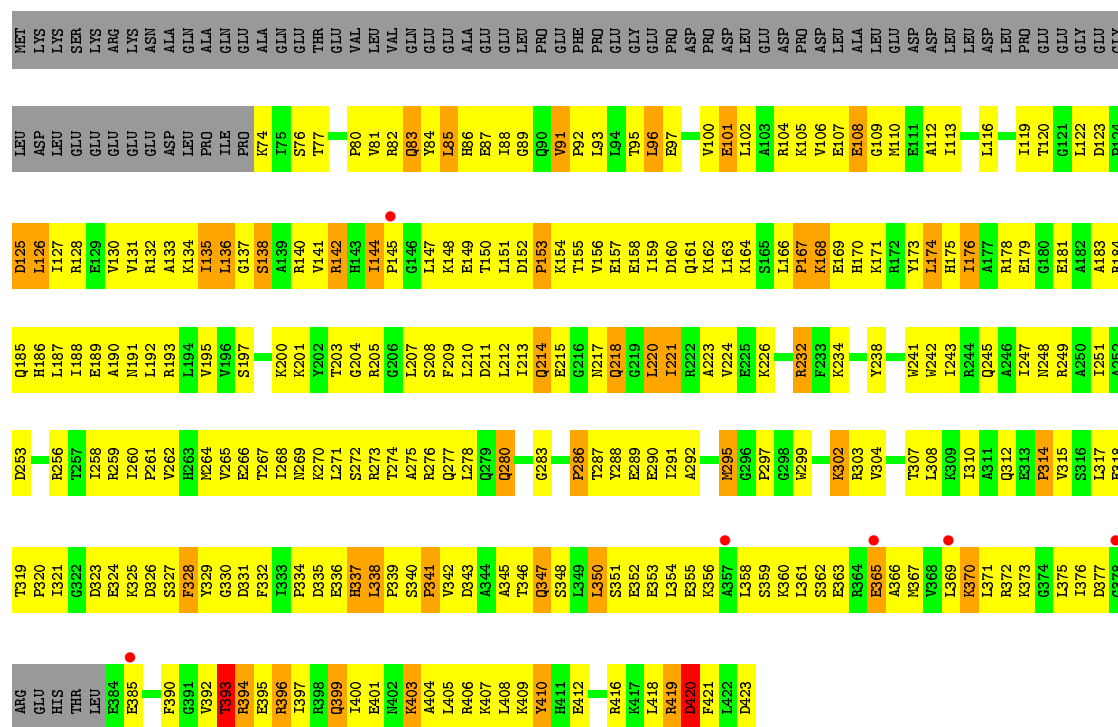
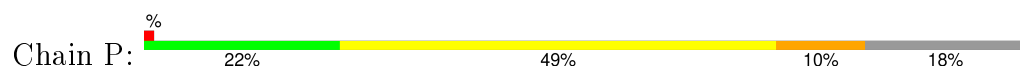


MET	K2	K3	E4	V5	R6	K7	V8	R9	I10	A11	L12	A13	S14	P15	E16	K17	I18	R19	S20	W21	Y22	G24	E25	V26	E27	K28	P29	E30	T31	I32	N33	R34	R35	T36	L37	K38	P39	E40	R41	D42	G43	L44	F45	D46	E47	R48	I49	F50	G51	P52	I53	K54	D55	Y56	E57	C58	A59	G60
-----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----





• Molecule 5: RNA polymerase sigma factor rpoD



4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	239.50Å 239.50Å 253.10Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	25.00 – 2.80 24.96 – 2.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (25.00-2.80) 92.0 (24.96-2.80)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.16 (at 2.80Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.231 , 0.268 0.226 , 0.261	Depositor DCC
R_{free} test set	21166 reflections (6.13%)	DCC
Wilson B-factor (Å ²)	58.9	Xtriage
Anisotropy	0.079	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 80.0	EDS
Estimated twinning fraction	0.499 for -h,-k,l 0.068 for h,-h-k,-l 0.068 for -k,-h,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.24$	Xtriage
Outliers	0 of 366401 reflections	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	58679	wwPDB-VP
Average B, all atoms (Å ²)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.92% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² 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	0.80	1/1838 (0.1%)	0.88	1/2498 (0.0%)
1	B	0.75	0/1838	0.82	2/2498 (0.1%)
1	K	0.75	0/1838	0.86	2/2498 (0.1%)
1	L	0.72	0/1838	0.78	1/2498 (0.0%)
2	C	0.84	0/8997	0.90	8/12164 (0.1%)
2	M	0.82	0/8997	0.89	7/12164 (0.1%)
3	D	0.84	0/10975	0.94	20/14836 (0.1%)
3	N	0.83	0/10975	0.93	18/14836 (0.1%)
4	E	0.84	0/783	0.97	0/1054
4	O	0.88	0/783	1.00	1/1054 (0.1%)
5	F	0.75	0/2812	0.82	3/3781 (0.1%)
5	P	0.75	0/2812	0.80	1/3781 (0.0%)
All	All	0.82	1/54486 (0.0%)	0.90	64/73662 (0.1%)

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
3	N	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	48	ILE	C-N	5.57	1.44	1.34

The worst 5 of 64 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	D	199	LEU	CA-CB-CG	-8.69	95.30	115.30
1	B	138	LEU	CA-CB-CG	8.01	133.72	115.30
3	N	1389	LEU	CA-CB-CG	7.77	133.18	115.30
5	P	136	LEU	CA-CB-CG	7.49	132.51	115.30
3	N	76	CYS	CA-CB-SG	6.73	126.11	114.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	N	132	TYR	Sidechain

5.2 Too-close contacts [i](#)

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	1806	0	1861	216	0
1	B	1806	0	1861	199	0
1	K	1806	0	1861	208	0
1	L	1806	0	1861	206	0
2	C	8829	0	8933	1184	0
2	M	8829	0	8933	1106	0
3	D	10797	0	10873	1450	0
3	N	10797	0	10873	1345	0
4	E	769	0	775	97	0
4	O	769	0	775	108	0
5	F	2771	0	2844	336	0
5	P	2771	0	2844	342	0
6	D	2	0	0	0	0
6	N	2	0	0	0	0
7	D	1	0	0	0	0
7	N	1	0	0	0	0
8	A	191	0	0	37	0
8	B	181	0	0	34	0
8	C	767	0	0	174	0
8	D	1100	0	0	234	0
8	E	93	0	0	14	0
8	F	333	0	0	58	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	K	151	0	0	30	0
8	L	179	0	0	49	0
8	M	739	0	0	195	0
8	N	1038	0	0	225	0
8	O	78	0	0	24	0
8	P	267	0	0	61	0
All	All	58679	0	54294	6401	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 59.

The worst 5 of 6401 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)
2:M:1054:THR:HG21	2:M:1079:PRO:HB3	1.28	1.11
1:A:95:GLN:HA	1:A:146:ARG:HH12	1.12	1.08
2:C:135:VAL:HG11	2:C:407:LYS:HA	1.36	1.04
3:N:1036:ARG:HH21	3:N:1042:ARG:HA	1.20	1.04
1:A:42:ARG:HH12	2:C:857:ASP:HB3	1.14	1.04

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	227/315 (72%)	202 (89%)	21 (9%)	4 (2%)	11	34
1	B	227/315 (72%)	202 (89%)	21 (9%)	4 (2%)	11	34
1	K	227/315 (72%)	200 (88%)	23 (10%)	4 (2%)	11	34
1	L	227/315 (72%)	204 (90%)	19 (8%)	4 (2%)	11	34
2	C	1117/1119 (100%)	917 (82%)	150 (13%)	50 (4%)	3	10

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	M	1117/1119 (100%)	907 (81%)	159 (14%)	51 (5%)	3	9
3	D	1388/1524 (91%)	1123 (81%)	191 (14%)	74 (5%)	2	7
3	N	1388/1524 (91%)	1110 (80%)	195 (14%)	83 (6%)	2	5
4	E	93/99 (94%)	74 (80%)	16 (17%)	3 (3%)	5	17
4	O	93/99 (94%)	75 (81%)	15 (16%)	3 (3%)	5	17
5	F	341/423 (81%)	283 (83%)	42 (12%)	16 (5%)	3	9
5	P	341/423 (81%)	285 (84%)	40 (12%)	16 (5%)	3	9
All	All	6786/7590 (89%)	5582 (82%)	892 (13%)	312 (5%)	3	9

5 of 312 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	29	GLU
1	B	29	GLU
1	B	48	ILE
2	C	152	PRO
2	C	178	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	202/273 (74%)	147 (73%)	55 (27%)	0	1
1	B	202/273 (74%)	163 (81%)	39 (19%)	2	5
1	K	202/273 (74%)	152 (75%)	50 (25%)	1	2
1	L	202/273 (74%)	158 (78%)	44 (22%)	1	3
2	C	941/941 (100%)	734 (78%)	207 (22%)	1	3
2	M	941/941 (100%)	730 (78%)	211 (22%)	1	3
3	D	1123/1279 (88%)	846 (75%)	277 (25%)	1	2
3	N	1123/1279 (88%)	866 (77%)	257 (23%)	1	3
4	E	83/87 (95%)	58 (70%)	25 (30%)	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
4	O	83/87 (95%)	64 (77%)	19 (23%)	1	3
5	F	295/370 (80%)	228 (77%)	67 (23%)	1	3
5	P	295/370 (80%)	249 (84%)	46 (16%)	3	9
All	All	5692/6446 (88%)	4395 (77%)	1297 (23%)	1	3

5 of 1297 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	E	30	LEU
1	L	25	LEU
3	N	1353	GLN
5	F	80	PRO
5	F	361	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 179 such sidechains are listed below:

Mol	Chain	Res	Type
4	E	86	GLN
1	L	212	ASN
3	N	1374	GLN
5	F	337	HIS
1	K	156	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 ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 6 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

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

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

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

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 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	229/315 (72%)	-0.37	0 100 100	38, 65, 92, 117	0
1	B	229/315 (72%)	-0.26	9 (3%) 43 31	53, 95, 115, 121	0
1	K	229/315 (72%)	-0.36	0 100 100	41, 66, 93, 122	0
1	L	229/315 (72%)	-0.38	3 (1%) 79 71	52, 94, 114, 127	0
2	C	1119/1119 (100%)	-0.34	8 (0%) 89 84	23, 81, 110, 119	0
2	M	1119/1119 (100%)	-0.38	6 (0%) 91 88	27, 79, 109, 122	0
3	D	1392/1524 (91%)	-0.32	12 (0%) 85 79	17, 68, 113, 130	0
3	N	1392/1524 (91%)	-0.34	16 (1%) 82 74	27, 69, 110, 138	0
4	E	95/99 (95%)	-0.34	0 100 100	47, 85, 115, 133	0
4	O	95/99 (95%)	-0.48	1 (1%) 82 74	40, 76, 97, 111	0
5	F	345/423 (81%)	-0.39	3 (0%) 85 79	46, 84, 112, 131	0
5	P	345/423 (81%)	-0.38	6 (1%) 73 63	54, 87, 112, 125	0
All	All	6818/7590 (89%)	-0.35	64 (0%) 85 79	17, 77, 111, 138	0

The worst 5 of 64 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	130	ALA	5.1
3	D	1240	THR	4.9
3	N	1243	THR	4.8
3	D	1243	THR	4.6
5	F	145	PRO	4.6

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	D	7112	1/1	0.99	0.14	0.51	75,75,75,75	0
6	ZN	N	7059	1/1	0.99	0.11	-0.82	100,100,100,100	0
7	MG	N	9002	1/1	0.94	0.13	-0.83	64,64,64,64	0
6	ZN	N	7113	1/1	0.98	0.10	-1.18	87,87,87,87	0
6	ZN	D	7058	1/1	0.98	0.07	-1.71	109,109,109,109	0
7	MG	D	9001	1/1	0.97	0.15	-	67,67,67,67	0

6.5 Other polymers [i](#)

There are no such residues in this entry.