



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 26, 2020 – 09:18 am BST

PDB ID : 5EXR  
Title : Crystal structure of human primosome  
Authors : Tahirov, T.H.; Baranovskiy, A.G.; Babayeva, N.D.  
Deposited on : 2015-11-24  
Resolution : 3.60 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

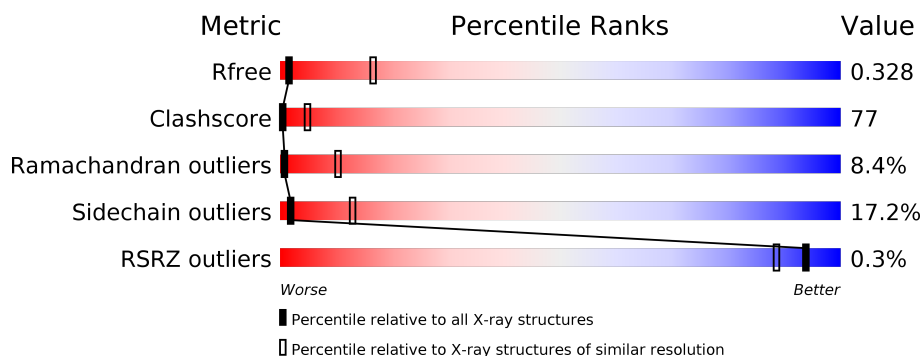
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.60 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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 respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	420	<div> <div>3%</div> <div>21% 63% 9% 7%</div> </div>
1	E	420	<div> <div>23% 62% 8% 7%</div> </div>
2	B	509	<div> <div>20% 48% 16% 15%</div> </div>
2	F	509	<div> <div>17% 50% 17% 15%</div> </div>
3	C	1128	<div> <div>21% 54% 18% 6%</div> </div>
3	G	1128	<div> <div>21% 53% 18% 6%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
4	D	597	
4	H	597	

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
6	SF4	B	601	-	-	X	-

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 37658 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 primase small subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	389	Total	C	N	O	S	0	0	0
			3261	2099	564	583	15			
1	E	389	Total	C	N	O	S	0	0	0
			3261	2099	564	583	15			

- Molecule 2 is a protein called DNA primase large subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	434	Total	C	N	O	S	0	0	0
			3562	2280	616	653	13			
2	F	434	Total	C	N	O	S	0	0	0
			3562	2280	616	653	13			

- Molecule 3 is a protein called DNA polymerase alpha catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	1057	Total	C	N	O	S	0	0	0
			8544	5477	1433	1578	56			
3	G	1057	Total	C	N	O	S	0	0	0
			8544	5477	1433	1578	56			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	516	ALA	VAL	engineered mutation	UNP P09884
G	516	ALA	VAL	engineered mutation	UNP P09884

- Molecule 4 is a protein called DNA polymerase alpha subunit B.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	444	Total	C	N	O	S	0	0	0
			3451	2194	576	666	15			

*Continued on next page...*

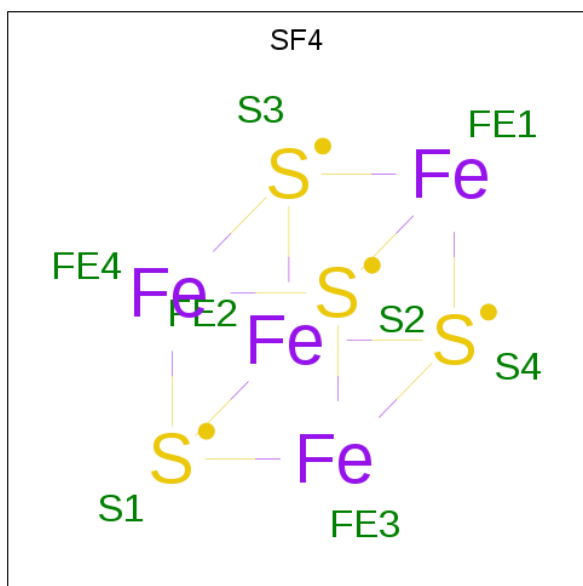
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	444	Total	C	N	O	S	0	0	0
			3451	2194	576	666	15			

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	G	2	Total	Zn	0	0
			2	2		
5	A	1	Total	Zn	0	0
			1	1		
5	C	2	Total	Zn	0	0
			2	2		
5	E	1	Total	Zn	0	0
			1	1		

- Molecule 6 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe<sub>4</sub>S<sub>4</sub>).

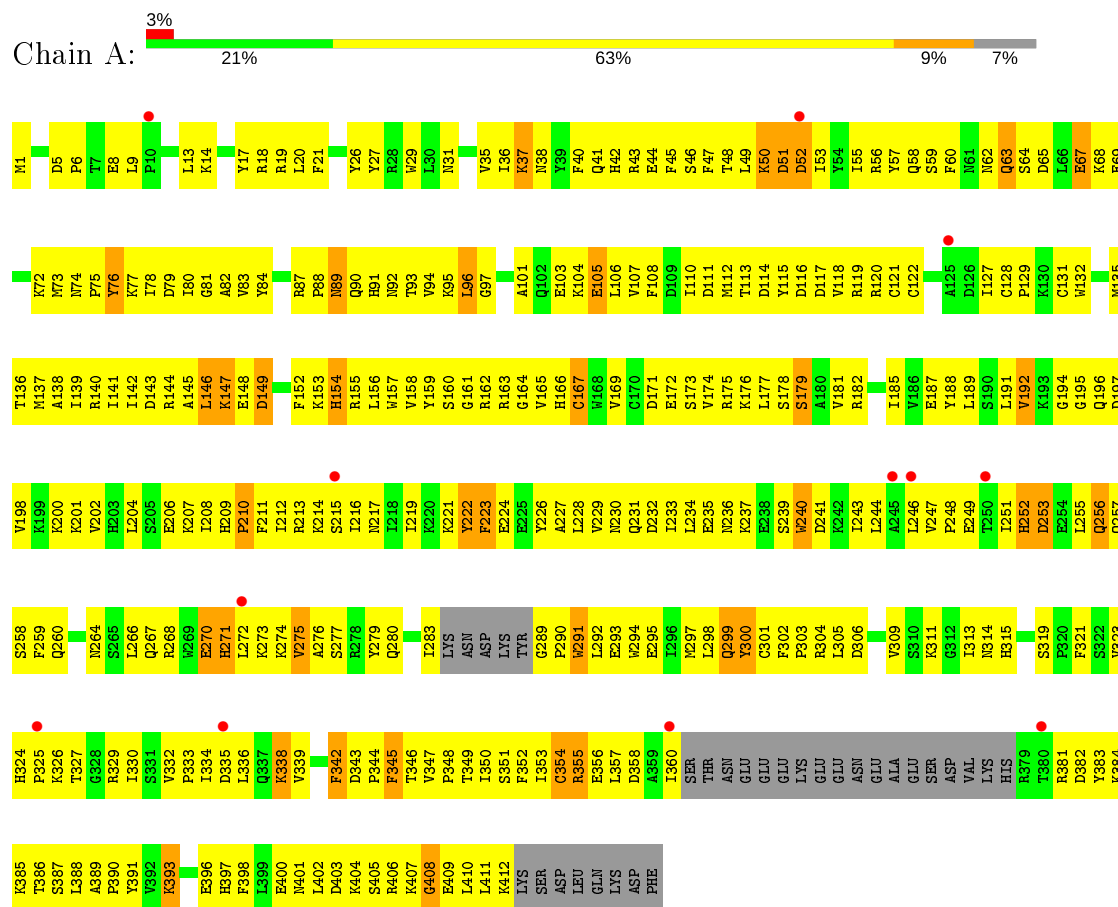


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	Fe	S	0	0
			8	4	4		
6	F	1	Total	Fe	S	0	0
			8	4	4		

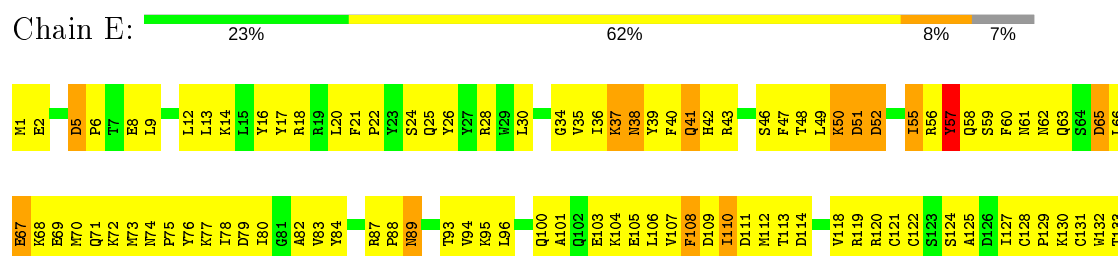
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

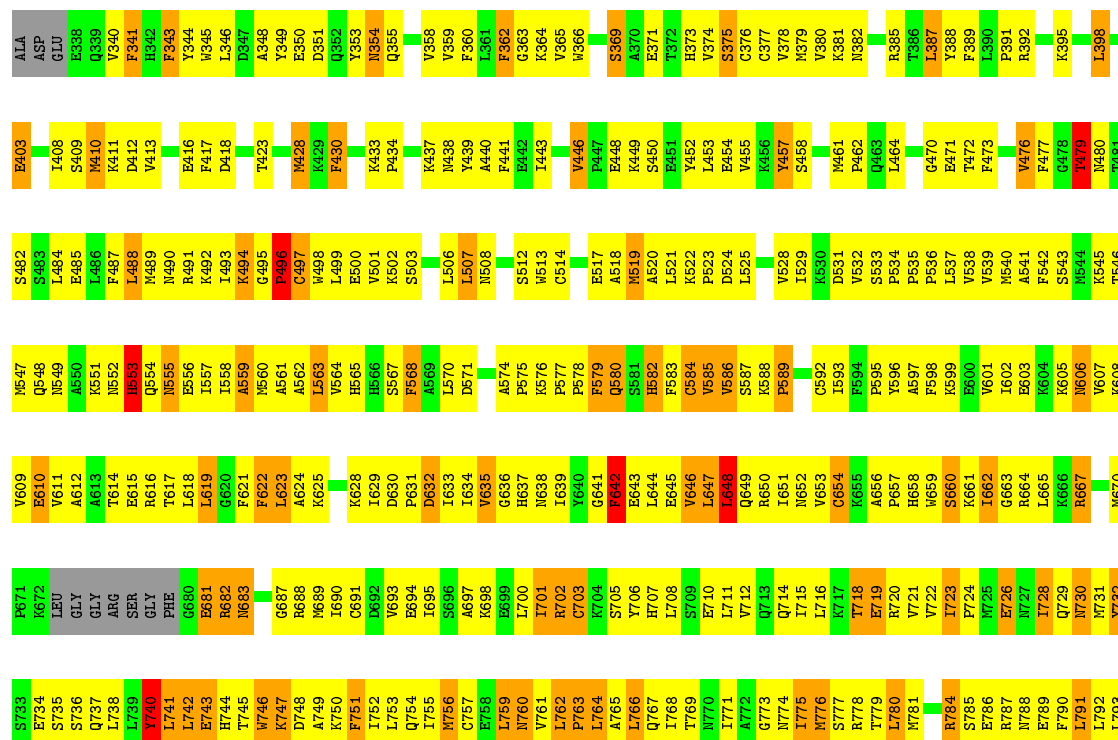
#### • Molecule 1: DNA primase small subunit



#### • Molecule 1: DNA primase small subunit

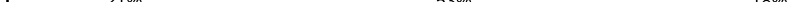








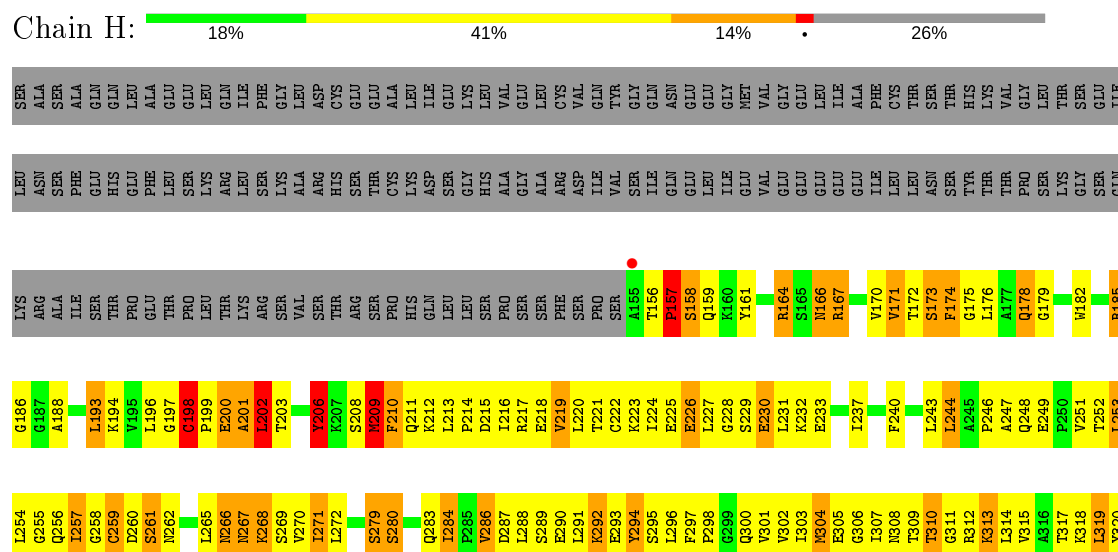
- Molecule 3: DNA polymerase alpha catalytic subunit

Chain G:  21% 53% 18% • 6%

F477	G478	T406	P407	I408	S409	M410	K411	D412	V413	E414	E415	E416	F417	D418	E419	K420	I421	Y425	K426	I427	M428	K429	F430	K433	K437	M438	Y439	A440	F441	W446	P447	E448	K449	Y452	L453	E454	V455	K456	Y457	S458	M461	P462	Q463	L464	L468	K469	G470	E471	L472	F473	S474	H475	V476																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
F477	G478	T406	P407	I408	S409	M410	K411	D412	V413	E414	E415	E416	F417	D418	E419	K420	I421	Y425	K426	I427	M428	K429	F430	K433	K437	M438	Y439	A440	F441	W446	P447	E448	K449	Y452	L453	E454	V455	K456	Y457	S458	M461	P462	Q463	L464	L468	K469	G470	E471	L472	F473	S474	H475	V476																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
G479	N480																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							

A1437	E1438	Q1439	F1440	L1441	S1442	R1443	S1444	G1445	Y1446	S1447	E1448	Y1449	N1450	L1451	L1454	F1455	A1456	GLY	CYS	ALA	VAL	LYS	SER																																													
M1375	K1376	A1377	T1378	L1379	Y1383	S1384	D1385	K1386	S1387	L1388	Y1389	T1390	Q1391	L1392	C1393	F1394	Y1395	R1396	Y1397	I1398	F1399	D1400	A1401	E1402	C1403	A1404	L1405	A1406	K1407	L1408	T1409	T1410	D1411	H1412	E1413	K1414	D1415	K1416	L1417	K1418	K1419	Q1420	F1421	F1422	T1423	P1424	V1426	L1427	Q1428	D1429	Y1430	R1431	K1432	L1433	K1434	L1435	T1436											
K1114	R1115	L1116	L1117	R1118	I1119	E1120	E1121	N1122	V1123	L1124	N1125	V1128	P1129	V1130	S1131	Q1132	F1133	E1134	L1135	N1136	L1139	T1140	K1141	D1142	P1143	Q1144	D1145	Y1146	P1147	D1148	K1149	S1150	S1151	H1154	V1155	H1156	V1157	A1158	L1159	H1160	I1161	N1162	S1163	V1175	S1176	Y1177	V1178	I1179	C1180	Q1181	D1182	G1183	N1184	M1185	F1246	R1247												
T1187	A1188	S1189	Q1190	R1191	A1192	Y1193	A1194	P1195	E1196	Q1197	L1198	Q1199	K1200	Q1201	D1202	P1203	L1204	F1205	I1206	D1207	T1208	Y1210	Y1211	Q1214	Q1215	I1216	H1217	P1218	V1219	V1220	A1221	C1222	I1223	C1224	E1225	P1226	I1227	D1228	G1229	I1230	D1231	A1232	V1233	I1235	A1236	T1237	H1238	M1239	E1240	E1241	E1242	G1298	S1299	G1300	T1301	D1302	M1303	E1304	P1305	D1243	T1244	Q1245	I1308	R1309				
V1248	H1249	H1250	K1251	H1252	K1253	D1254	E1255	E1256	N1257	D1258	ALA	LEU	GLY	PRQ	ALA	Q1266	L1267	T1268	E1271	K1272	Y1273	R1274	D1275	C1276	P1277	R1278	C1283	P1284	T1285	C1286	N1290	Y1291	Y1292	K1293	D1294	V1295	F1296	D1297	G1298	S1299	G1300	T1301	D1302	M1303	E1304	P1305	D1243	T1244	Q1245	I1308	R1309																	
I1313	D1314	C1315	K1316	A1317	S1318	P1319	L1320	T1321	F1322	Q1325	L1326	S1327	M1328	K1329	L1330	I1331	M1332	D1333	R1334	R1335	F1337	L1338	K1339	K1340	Y1341	L1342	C1348	E1349	E1350	P1351	T1352	R1353	L1354	M1355	R1356	T1357	R1358	H1359	L1360	Q1363	F1364	S1365	R1366	T1367	G1368	P1369	L1370	C1371	P1372	A1373	C1374																	
M1375	K1376	A1377	T1378	L1379	Y1383	S1384	D1385	K1386	S1387	L1388	Y1389	T1390	Q1391	L1392	C1393	F1394	Y1395	R1396	Y1397	I1398	F1399	D1400	A1401	E1402	C1403	A1404	L1405	A1406	K1407	L1408	T1409	T1410	D1411	H1412	E1413	K1414	D1415	K1416	L1417	K1418	K1419	Q1420	F1421	F1422	T1423	P1424	V1426	L1427	Q1428	D1429	Y1430	R1431	K1432	L1433	K1434	L1435	T1436											
H794	A795	F796	Y797	E798	N799	H800	Y801	R802	S803	L804	Y805	P806	Q807	R808	PHE	D748	ARG	A749	K750	V751	I752	L753	Q754	I755	V756	D630	H631	D632	S667	F668	A669	L570	P575	L638	I639	Y640	Q641	F642	H643	L644	E645	V646	L647	S686	S687	K688	L649	Q649	K650	P651	N652	V653	C654	K655	P656	Y657	H658	F598	K599	E600	K661	I662	V601	G663	R664	L665	K666	R667
K854	F855	R856	L857	L858	L859	D860	F861	N862	S863	L864	Y865	P866	Q867	R868	I869	ARG	A749	K750	V751	I752	L753	Q754	I755	V756	D630	H631	D632	S667	F668	A669	L570	P575	L638	I639	Y640	Q641	F642	H643	L644	E645	V646	L647	S686	S687	K688	L649	Q649	K650	P651	N652	V653	C654	K655	P656	Y657	H658	F598	K599	E600	K661	I662	V601	G663	R664	L665	K666	R667	
E914	I915	R916	K917	L918	V919	E920	R921	N922	V925	L926	Q927	R928	D933	L934	N935	T936	ARG	A749	K750	V751	I752	L753	Q754	I755	V756	D630	H631	D632	S667	F668	A669	L570	P575	L638	I639	Y640	Q641	F642	H643	L644	E645	V646	L647	S686	S687	K688	L649	Q649	K650	P651	N652	V653	C654	K655	P656	Y657	H658	F598	K599	E600	K661	I662	V601	G663	R664	L665	K666	R667
R981	E982	R983	L984	N985	H986	T987	R988	E989	N990	V991	Q992	K993	N995	L996	E997	S998	ARG	A749	K750	V751	I752	L753	Q754	I755	V756	D630	H631	D632	S667	F668	A669	L570	P575	L638	I639	Y640	Q641	F642	H643	L644	E645	V646	L647	S686	S687	K688	L649	Q649	K650	P651	N652	V653	C654	K655	P656	Y657	H658	F598	K599	E600	K661	I662	V601	G663	R664	L665	K666	R667
F1044	K1045	S1046	L1047	L1048	L1049	L1050	K1051	Y1055	A1056	A1057	L1058	V1059	V1060	T1063	S1064	Y1068	L1074	K1075	G1076	L1077	D1078	R1081	Q1082	P1083	L1084	D1085	A1086	K1089	D1090	T1091	G1092	M1093	F1094	V1095	X1096	G1097	Q1098	I1099	L1100	S1101	D1102	Q1103	S1104	R1105	L1106	T1107	C1108	Y1109	E1110	N1111	T1112	Q1113																
K1114	R1115	L1116	L1117	R1118	I1119	E1120	E1121	N1122	V1123	L1124	N1125	V1128	P1129	V1130	S1131	Q1132	F1133	E1134	L1135	N1136	L1139	T1140	K1141	D1142	P1143	Q1144	D1145	Y1146	P1147	D1148	K1149	S1150	S1151	H1154	V1155	H1156	V1157	A1158	L1159	H1160	I1161	N1162	S1163	V1175	S1176	Y1177	V1178	I1179	C1180	Q1181	D1182	G1183	N1184	M1185	F1246	R1247												
T1187	A1188	S1189	Q1190	R1191	A1192	Y1193	A1194	P1195	E1196	Q1197	L1198	Q1199	K1200	Q1201	D1202	P1203	L1204	F1205	I1206	D1207	T1208	Y1210	Y1211	Q1214	Q1215	I1216	H1217	P1218	V1219	V1220	A1221	C1222	I1223	C1224	E1225	P1226	I1227	D1228	G1229	I1230	D1231	A1232	V1233	I1235	A1236	T1237	H1238	M1239	E1240	E1241	E1242	G1298	S1299	G1300	T1301	D1302	M1303	E1304	P1305	D1243	T1244	Q1245	I1308	R1309				
V1248	H1249	H1250	K1251	H1252	K1253	D1254	E1255	E1256	N1257	D1258	ALA	LEU	GLY	PRQ	ALA	Q1266	L1267	T1268	E1271	K1272	Y1273	R1274	D1275	C1276	P1277	R1278	C1283	P1284	T1285	C1286	N1290	Y1291	Y1292	K1293	D1294	V1295	F1296	D1297	G1298	S1299	G1300	T1301	D1302	M1303	E1304	P1305	D1243	T1244	Q1245	I1308	R1309																	
I1313	D1314	C1315	K1316	A1317	S1318	P1319	L1320	T1321	F1322	Q1325	L1326	S1327	M1328	K1329	L1330	I1331	M1332	D1333	R1334	R1335	F1337	L1338	K1339	K1340	Y1341	L1342	C1348	E1349	E1350	P1351	T1352	R1353	L1354	M1355	R1356	T1357	R1358	H1359	L1360	Q1363	F1364	S1365	R1366	T1367	G1368	P1369	L1370	C1371	P1372	A1373	C1374																	
M1375	K1376	A1377	T1378	L1379	Y1383	S1384	D1385	K1386	S1387	L1388	Y1389	T1390	Q1391	L1392	C1393	F1394	Y1395	R1396	Y1397	I1398	F1399	D1400	A1401	E1402	C1403	A1404	L1405	A1406	K1407	L1408	T1409	T1410	D1411	H1412	E1413	K1414	D1415	K1416	L1417	K1418	K1419	Q1420	F1421	F1422	T1423	P1424	V1426	L1427	Q1428	D1429	Y1430	R1431	K1432	L1433	K1434	L1435	T1436											

● Molecule 4: DNA polymerase alpha subunit B





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	113.10Å 210.16Å 172.56Å 90.00° 93.56° 90.00°	Depositor
Resolution (Å)	39.94 – 3.60 49.73 – 3.30	Depositor EDS
% Data completeness (in resolution range)	68.9 (39.94-3.60) 73.5 (49.73-3.30)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 3.33Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.268 , 0.326 0.275 , 0.328	Depositor DCC
$R_{free}$ test set	4621 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	78.1	Xtriage
Anisotropy	0.047	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 69.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.24$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.86	EDS
Total number of atoms	37658	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	61.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 44.59 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.4935e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, SF4

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	0.51	0/3343	0.72	0/4508
1	E	0.48	0/3343	0.67	0/4508
2	B	0.57	0/3646	0.82	5/4908 (0.1%)
2	F	0.57	0/3646	0.80	3/4908 (0.1%)
3	C	0.58	0/8724	0.83	3/11788 (0.0%)
3	G	0.58	1/8724 (0.0%)	0.83	5/11788 (0.0%)
4	D	0.61	0/3529	0.86	2/4795 (0.0%)
4	H	0.60	0/3529	0.86	3/4795 (0.1%)
All	All	0.57	1/38484 (0.0%)	0.81	21/51998 (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	B	0	2
2	F	0	1
3	C	0	1
3	G	0	1
4	D	0	1
4	H	0	1
All	All	0	7

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	1371	CYS	CB-SG	-5.25	1.73	1.81

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	368	LEU	CA-CB-CG	-7.73	97.53	115.30
2	B	368	LEU	CA-CB-CG	-7.12	98.92	115.30
4	D	202	LEU	CA-CB-CG	6.87	131.10	115.30
4	H	202	LEU	CA-CB-CG	6.48	130.21	115.30
3	C	742	LEU	CA-CB-CG	-6.32	100.76	115.30

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	110	TYR	Sidechain
2	B	309	TYR	Sidechain
3	C	740	TYR	Sidechain
4	D	349	TYR	Sidechain
2	F	110	TYR	Sidechain

## 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	3261	0	3247	444	0
1	E	3261	0	3247	419	0
2	B	3562	0	3542	548	0
2	F	3562	0	3542	557	0
3	C	8544	0	8632	1431	0
3	G	8544	0	8634	1426	0
4	D	3451	0	3425	535	0
4	H	3451	0	3425	532	0
5	A	1	0	0	0	0
5	C	2	0	0	0	0
5	E	1	0	0	0	0
5	G	2	0	0	0	0
6	B	8	0	0	3	0
6	F	8	0	0	0	0
All	All	37658	0	37694	5765	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 77.

The worst 5 of 5765 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:G:858:LEU:HD13	3:G:1007:MET:HG3	1.23	1.21
2:B:209:VAL:HG12	2:B:210:PRO:HD2	1.25	1.19
3:C:730:ASN:ND2	3:C:730:ASN:H	1.28	1.19
4:D:476:LEU:HD11	4:D:502:ILE:HD11	1.22	1.14
1:E:20:LEU:HD21	1:E:357:LEU:HD22	1.16	1.13

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	383/420 (91%)	326 (85%)	45 (12%)	12 (3%)	4	32
1	E	383/420 (91%)	336 (88%)	36 (9%)	11 (3%)	4	33
2	B	432/509 (85%)	291 (67%)	93 (22%)	48 (11%)	0	6
2	F	432/509 (85%)	295 (68%)	84 (19%)	53 (12%)	0	5
3	C	1047/1128 (93%)	731 (70%)	226 (22%)	90 (9%)	1	10
3	G	1047/1128 (93%)	743 (71%)	209 (20%)	95 (9%)	1	9
4	D	442/597 (74%)	325 (74%)	79 (18%)	38 (9%)	1	10
4	H	442/597 (74%)	330 (75%)	73 (16%)	39 (9%)	1	9
All	All	4608/5308 (87%)	3377 (73%)	845 (18%)	386 (8%)	1	10

5 of 386 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	37	LYS
2	B	29	TYR
2	B	35	GLU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	B	90	GLU
2	B	94	GLU

### 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	363/393 (92%)	322 (89%)	41 (11%)	6	30
1	E	363/393 (92%)	324 (89%)	39 (11%)	6	32
2	B	394/459 (86%)	326 (83%)	68 (17%)	2	12
2	F	394/459 (86%)	329 (84%)	65 (16%)	2	15
3	C	962/1013 (95%)	785 (82%)	177 (18%)	1	10
3	G	962/1013 (95%)	780 (81%)	182 (19%)	1	9
4	D	390/526 (74%)	314 (80%)	76 (20%)	1	9
4	H	390/526 (74%)	312 (80%)	78 (20%)	1	8
All	All	4218/4782 (88%)	3492 (83%)	726 (17%)	2	13

5 of 726 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	D	414	ARG
2	F	76	SER
4	H	266	ASN
4	D	477	LEU
1	E	51	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 180 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	1380	GLN
1	E	42	HIS
3	G	1420	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	D	166	ASN
4	D	419	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 8 ligands modelled in this entry, 6 are monoatomic - leaving 2 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	$\# Z  > 2$
6	SF4	F	601	2	0,12,12	0.00	-	-		
6	SF4	B	601	2	0,12,12	0.00	-	-		

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
6	SF4	F	601	2	-	-	0/6/5/5
6	SF4	B	601	2	-	-	0/6/5/5

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.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	601	SF4	3	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 ⓘ

### 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	389/420 (92%)	-0.09	12 (3%) 49 33	39, 101, 121, 137	0
1	E	389/420 (92%)	-0.40	1 (0%) 94 88	45, 98, 115, 136	0
2	B	434/509 (85%)	-0.72	1 (0%) 95 91	5, 61, 112, 135	0
2	F	434/509 (85%)	-0.66	0 100 100	4, 64, 115, 135	0
3	C	1057/1128 (93%)	-0.75	0 100 100	1, 51, 93, 123	0
3	G	1057/1128 (93%)	-0.73	0 100 100	2, 54, 96, 116	0
4	D	444/597 (74%)	-0.70	0 100 100	1, 44, 94, 111	0
4	H	444/597 (74%)	-0.68	1 (0%) 95 91	2, 47, 94, 121	0
All	All	4648/5308 (87%)	-0.64	15 (0%) 94 88	1, 60, 109, 137	0

The worst 5 of 15 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	125	ALA	3.6
1	A	360	ILE	2.9
1	A	272	LEU	2.9
1	A	215	SER	2.7
1	A	245	ALA	2.7

### 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 [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 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	B-factors( $\text{\AA}^2$ )	Q<0.9
5	ZN	A	501	1/1	0.91	0.04	123,123,123,123	0
5	ZN	E	501	1/1	0.94	0.10	91,91,91,91	0
5	ZN	G	1501	1/1	0.95	0.16	78,78,78,78	0
6	SF4	F	601	8/8	0.97	0.18	1,1,8,18	0
5	ZN	C	1502	1/1	0.97	0.15	26,26,26,26	0
6	SF4	B	601	8/8	0.98	0.18	1,1,2,9	0
5	ZN	G	1502	1/1	0.99	0.12	1,1,1,1	0
5	ZN	C	1501	1/1	0.99	0.14	46,46,46,46	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.