

wwPDB X-ray Structure Validation Summary Report (i)

Apr 17, 2024 – 12:36 PM EDT

PDB ID	:	8UKQ
Title	:	RNA polymerase II elongation complex with Fapy-dG lesion in apo state
Authors	:	Hou, P.; Oh, J.; Wang, D.
Deposited on	:	2023-10-15
Resolution	:	3.50 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 https://www.wwpdb.org/validation/2017/XrayValidationReportHelp with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) 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	:	2.36.1
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)	:	2.36.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure: $X\text{-}RAY \, DIFFRACTION$

The reported resolution of this entry is 3.50 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	$egin{array}{c} { m Whole \ archive} \ (\#{ m Entries}) \end{array}$	${f Similar\ resolution}\ (\#{ m Entries,\ resolution\ range}({ m \AA}))$
R _{free}	130704	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)
RNA backbone	3102	1002 (4.00-3.00)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	G	Quality of chain	
1	R	9	44%	44%	11%
2	Т	29	17%	66%	17%
3	Ν	18	28%	44%	28%
4	А	1733	3% 46%	32%	• 20%

Continued on next page...



Continued from previous page...

Mol	Chain	Length	Quali	ty of chain
5	В	1224	.% 5 5%	36% • 8%
6	С	318	3% 52%	31% · 16%
7	Е	215	9%	41%
8	F	155	.%	6 % • 45%
9	Н	146	53%	36% • 9%
10	Ι	122	3% 57%	36% • •
11	J	70	50%	40% · 7%
12	Κ	120	^{2%} 52%	43% 5%
13	L	70	<u>6%</u> <u>36%</u>	26% 39%



2 Entry composition (i)

There are 15 unique types of molecules in this entry. The entry contains 28981 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 RNA chain called RNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	9	Total 194	C 88	N 40	O 58	Р 8	0	0	0

• Molecule 2 is a DNA chain called tsDNA with FapydG lesion.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
2	Т	24	Total 481	C 230	N 76	0 151	Р 24	0	0	0

• Molecule 3 is a DNA chain called ntsDNA.

Mol	Chain	Residues		Ate	oms			ZeroOcc	AltConf	Trace
3	Ν	13	Total 275	C 128	N 61	O 73	Р 13	0	0	0

• Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues		A	toms		ZeroOcc	AltConf	Trace	
4	А	1384	Total 10828	C 6831	N 1896	O 2041	S 60	0	0	0

• Molecule 5 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues		Α	toms		ZeroOcc	AltConf	Trace	
5	В	1126	Total 8874	C 5616	N 1555	O 1650	S 53	0	0	0

• Molecule 6 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues		At	oms		ZeroOcc	AltConf	Trace	
6	С	267	Total 2101	C 1320	N 349	0 419	S 13	0	0	0



• Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues		At	oms		ZeroOcc	AltConf	Trace	
7	Е	212	Total 1731	C 1100	N 305	0 315	S 11	0	0	0

• Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
8	F	86	Total 684	C 437	N 115	0 129	${ m S} { m 3}$	0	0	0

• Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
9	Н	133	Total 1064	C 670	N 179	O 211	$\frac{S}{4}$	0	0	0

• Molecule 10 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues		\mathbf{A}^{\dagger}	toms			ZeroOcc	AltConf	Trace
10	Ι	118	Total 952	C 585	N 173	0 184	S 10	0	0	0

• Molecule 11 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues		Atc	\mathbf{ms}			ZeroOcc	AltConf	Trace
11	J	65	Total 532	C 339	N 93	0 94	S 6	0	0	0

• Molecule 12 is a protein called DNA-directed RNA polymerase II subunit RPB11.

Mol	Chain	Residues		At	oms			ZeroOcc	AltConf	Trace
12	K	114	Total 919	C 590	N 156	0 171	${ m S} { m 2}$	0	0	0

• Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues		Ato	\mathbf{ms}			ZeroOcc	AltConf	Trace
13	L	43	Total 337	C 208	N 66	O 59	$\frac{S}{4}$	0	0	0

• Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
14	А	2	Total Zn 2 2	0	0
14	В	1	Total Zn 1 1	0	0
14	С	1	Total Zn 1 1	0	0
14	Ι	2	Total Zn 2 2	0	0
14	J	1	Total Zn 1 1	0	0
14	L	1	Total Zn 1 1	0	0

• Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	А	1	Total Mg 1 1	0	0



3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: RNA

N306 D307	1308 1	0311 P312	4313 0 0 0 0	B320	P321	K323	1325 1325	R326 A327	R328	K332	1336	R337 G338	N339 1 340	M341	G342 K343	R344	R350	T351 V352	I353	D356	E360 L361	D362 D363	K368	5369 1370		1375 Y376	E377 E378	V380	R387 1 388	L391 V392
H399	L415	R420	D423	1124 0425 1.426		K431	E433	R434 H435	1436	N439	V442	L443	R446	P448	H451	K452	M455	M456 A457	H458 R459	1463	P464 Y465	S466 T467	F468 R469	L470 N471	L472	5473 V474	A480	D483 C484	LOTD A	H490
R498 A499	E500 L501	A506	P508	V510	S513	6115 0515 8516	N517	K518 P519	C520 M521	G522	0701	D526 T527	L528	G530	1531 R532	K533 1.534	1535 T535	L536 R537	D538 T539	F540	V5 <mark>46</mark> L547	N548 M549	L550 V551			V 559	P563 A564	1566 1566 8567	P568 KF60	P570 L571 <mark>W572</mark>
S573 G574	K575 Q576	1577 L578		P583	H587	D592	T595	T596 L597	MGO5	L606	1608	0611	1612 1613	F614	V617	E618 K619	K620	G623	S624 S625	N626	V6 <mark>33</mark> T634	R635	L645	N648	K651	V652 V653	W656		CA65	1666 16667 16687
T669 1670	<mark>A671</mark> D672	G673 P674	16/5 M676 D677	E678		4684	K687	K 688	L691	K695	<mark>0698</mark>	A704	1 710	R711	V718	6621	77	D727 K728	A729 G730	R731 L732	N736	1.740	N741 N742	V743 K744	Q745	K752	1756	Q760 M761		6766 6766
R774 1775	<mark>A776</mark> F777	G778 F779	V / 80 D781 D760	T783	P7 85		D7 90	P794	E7 95 S7 96	K797	F7 99	V800 E801	N802 5803	Y804	L805 R806	FR10	F813	F814 F815	M818	G819 G820	L824	V836	R830	R840	L845	E846 D847	H851	T856 B857	N858 S850	L860 C861 N862
0865	F866 1867	E870	08/1 G872 M672	D874 D875	A876 1027	1878 1878	K880	0881 8882	D890		0600 N899	D900 L901	L902 Mana		H906 T907	L908		L912 L913	E914 S915	G916	1919	L929	Е932 У 933	936		D939 R940	K941 F942	L943 R944 F045	N946	E951 A952 N953
W954 P955	L956	1960 R961	1072 1072	1975			DOGT	066A	L993 0994	E995	D G G M	L1000	N1004	I1006		A1010	R1012	D1013 A1014	V1015	F1018	L1021 L1022	R1023 S1024	R1025	A1027 T1028	R1029	K1030 V1031	L1032 Q1033	E1034 Y1035 B1036	L1037	F1042
D1043 W1044	<mark>V1045</mark> L1046	S1047 N1048	11049 E1050 A1054	R1055			000Tp	A1069 Q1070	S1071	L1081	THR	PHE HIS	PHE AT A	GLY	VAL ALA	SER K1092	K1093	P1099	E1103	I1104 L1105	M1111	P1114	S1115	V1118	01128	E1129 Q1130	A1131 K1132	L1133 11134 B1135	S1136	11138 E1139 H1140
L1143	I1148	A1149 S1150		E1105	D1166	E1168		Q1171 L1172	L1176	LEU	GLU	GLU ALA	0TD	SER	ASP	GLN D1188		L1192 L1193	R1194 L1195	E1196 L1197	D1198 R1199	CUC IM	D1 206		V1212	61213 E1214	R1215 11216	K1217 Q1218 T1319	F1220 F1220	L1224 F1225
V1226 11227	<mark>W1228</mark> S1229	E1234	K1235 L1236 T1027	11237 11238 R1239	C1240	V1242	ARG	PRO LYS	SER LEU	ASP	GLU	GLU	ALA	GLU	D1257	K1262 T1263	E1264	M1 267	L1268	11271 T1272	L1273 R1274	E1 977	N1278	V1282 V1283	M1284	M1285 K1286	Y1287 D1288	K1289 K1290 V1201	P1292	T1295 G1296 E1297
K1300	E1301 P1302	E1303 W1304	11306	T1308	N1312	M1317		T1325	Y1328	N1330	F1332	11333 D1334	11335 M1336	E1337	11341	E1342	R1345	A1346 A1347	L1348 Y1349	K1350	N1 <mark>364</mark> Y1365	R1366	L1371 V1372	D1373	T1376	113//	T1382 S1383	V1384 T1385 B1386		L1397
M1398 R1399	F1402	E1403 E1404	11405 V1406 E1167	11408	E1417 11110		C1421 R1422	S1425	E1426 N1427	V1428	L1430	G1431	A1434 D1435	I1436	G1437 T1438	F1441		D1446 GLU	GLU SER	LEU VAL	LYS TYR	MET	GLU	LYS	THR	GLU	GLU ASP	den Asp	GLY	VAL THR PRO
TYR SER	ASN GLU	SER GLY	VAL	ALA	LEU	VAL	ASP	GLU	MET PHE	SER	LEU	VAL ASP	SER	SER	ASP	ALA	ALA	GLY	THR	ALA TYR	GLY GLY	ALA	TYR	GLU	E	PRO	PHE GLY	TYR CI V	GLU AT A	PRO THR SER
PR0 GLY	PHE GLY	VAL SER	PRO PRO	PHE	PRO	SER	THR	TYR SER	PRO THR	SER	ALA	TYR SER	PRO THR	SER	PRU SER	TYR	PRO	THR SER	PRO SER	TYR SER	PR0 THR	SER	SER	SER	THR	PRO	TYR	PRO THR	SER	SER TYR SER



• Molecule 5: DNA-directed RNA polymerase II subunit RPB2

Cha	in	B:	.%								[55%	6														3	6%						•	8%				
MET SER ASP	LEU ALA	ASN	GLU	LYS TYR	TYR	GLU	ASP	PKU TYR	GLY	PHE	GLU GLU		A23	P24 T25	071	D29	830	V33	134	835 136	A30 F37	F38	R39 E40		V44	S50	F51	V55	163	5	768 T68	L69	0/1	A75	SIH	THR	GLU	SER ASP	ASN
ILE SER R86	K87 Y88	E89	S91	F92	I95	K99	P100	M101 V102	N103	E104	H110	A111	L112	Y113	7 773	A117	R118	LII9 R120	N121	L122	S126	G127	L128	K133	E138	ALA	ILE ASP	VAL	GLY	ARG	LEU	TAS	GLU	LEU	ALA	GLU	SER	GLU ASP	ASP
SER GLU <mark>S162</mark>	F166	1167 1168		L181 S182	E183	E186		L189	E194	C195	P196 F197	D198	M1 99	2000		V211	L212 T042	5121	N221	1222 1223	0224 0224		K228	1234	R241	S242	A243 L244		K249	Q255	V250 K257	L258	R261	E262	6263 S264	20C	T268	1269 K270	A271
T272 L273	1276	D279 1280	P281	1282 V283	1284 TOBE	1200 F286	R287	1292	P293		E296 T007	L298	E299	H300	L311		L314 V215	P316	C317		5221 F322		R327	F333	I334	R337	GLY THR	ALA	GLY	ILE	LIS K345	E346	1349	Q350	K353	D354	K358	E359 F360	L361
13 <mark>64</mark>	8372 R373	K374	F376	F377 L378	G379	M381	1382	N383 R384	L385	L386	L387	L390	D391	R392 V303	D394	Q 395	D396	H400		K404	L406	D407	L408	A414	Q415 L416	F417	K418	L424	1425 K426		r430 Y431		E438 ALA	SIH	ASP PHE	ASN	SAT	L446 A447	1448
L457	L461	G464 N465	W466	K470	C L VN	5474 S474		S480 0481	V482	L483	N484 R485	Y486	T487	Y488	L492	-	L495	R490 R497		1502	ARG	ASP	GLY LYS	LEU	A509	R512	Q513 L514	H515		W519	4020 L521		A525 E526	T527	A532	C533	100	N538	C544
1545 S546	P551	I555	M563	L566	E567	V570		d573 ●	A577	T578	K579 V580	F581		G584 VE85	W586	H587	G588	N592		R595	E598	-	R601 T602	L603	R604 R605	-	E612 V613	S614	M615 I616	R617	1619 I619	R620	E621 K622	E623	F627	2 2 2 2 2	R632	V633 Y634	R635
P636 L637 F638	1639 V640	E641 D642	D643	L646	G647	по <u>4</u> 0 К649	E650	L651 K652	V653	R654	TEE	A659	K660	L661	T664	E665	Y666	D668	ILE	0TD	GLY	PHE	GLU ASP	V676	Y679	T680	W681	<mark>G688</mark>	L689 V690	E691	I692 I693		F697	1701	<mark>q706</mark>	P7.07 E7.08	D7 09	L710	E7 14
ALA ASN GLU	GLU ASN	ASP 1721		K/28 1729	CC 2-11	H/ 33	A735	T736 T737	F738	T739	H740	E742		1748 1740		A753	S754	1756 1756	P757			<mark>q763</mark>	S764 P765	R766	S771		K775 0776	A777	M/ /8 G779	V780	r / 81 L782		1795 L796		0800 0800	K801 D800	L803	G804 T805	T806
R807 A808 M809	E810 Y811	L812	R815	E816 L817		NOZZ	V825	A828	C829	Y830	M834	0 835	E836	D837	I840	M841	N842	u 843 S844	S845	1846 De47	L04 / R848	G849	L850 F851		F856 R857	S858	Y859	Q862	F023	Y866	4007 M868	S869	E872	T873	F8/4 E875	K876 D877	ror r 0878	N881	T882
L883 R884 M885	K886	K892 I 803	D894	D896 D896	G897	1899 1899	A900	N903	R904	V905		V910		K914 T015	T916		S919	ASP	GLU	0TD	TEU	GLY	GLN ARG	THR	A930 Y931	H932	S933 K934	R935	D936 A937	<mark>8938</mark>	R942		F345	1948	V949 D950	0951 10650	L953	V954 T955	T956







4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants	161.54Å 223.49Å 191.47Å	Depositor
a, b, c, α , β , γ	90.00° 98.38° 90.00°	Depositor
Bosolution (Å)	48.44 - 3.50	Depositor
	48.44 - 3.50	EDS
% Data completeness	$99.4 \ (48.44 - 3.50)$	Depositor
(in resolution range)	$99.4 \ (48.44 - 3.50)$	EDS
R_{merge}	0.55	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) > 1$	$1.19 (at 3.48 \text{\AA})$	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
B B.	0.268 , 0.302	Depositor
Π, Π_{free}	0.271 , 0.305	DCC
R_{free} test set	1958 reflections (2.33%)	wwPDB-VP
Wilson B-factor $(Å^2)$	94.4	Xtriage
Anisotropy	0.550	Xtriage
Bulk solvent $k_{sol}(e/Å^3), B_{sol}(Å^2)$	0.26 , 84.2	EDS
L-test for $twinning^2$	$ < L >=0.37, < L^2>=0.20$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	28981	wwPDB-VP
Average B, all atoms $(Å^2)$	123.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 2.71% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities estimated from amplitudes.

5 Model quality (i)

5.1 Standard geometry (i)

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

Mal	Chain	Bond	lengths	Bo	ond angles
MOI	Ullalli	RMSZ	# Z > 5	RMSZ	# Z > 5
1	R	0.29	0/218	0.94	0/339
2	Т	0.58	0/507	1.05	0/775
3	N	0.54	0/311	0.77	0/479
4	А	0.27	0/11020	0.55	1/14907~(0.0%)
5	В	0.27	0/9046	0.53	0/12210
6	С	0.28	0/2139	0.50	0/2899
7	Е	0.28	0/1767	0.52	0/2378
8	F	0.26	0/696	0.54	0/943
9	Н	0.30	0/1082	0.65	2/1466~(0.1%)
10	Ι	0.28	0/970	0.57	0/1308
11	J	0.27	0/541	0.62	0/727
12	K	0.29	0/937	0.55	0/1265
13	L	0.27	0/339	0.62	0/450
All	All	0.29	0/29573	0.57	3/40146~(0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
9	Н	92	ASP	CB-CG-OD1	5.56	123.30	118.30
4	А	472	LEU	CB-CG-CD2	-5.36	101.89	111.00
9	Н	125	LEU	CA-CB-CG	5.30	127.49	115.30

There are no chirality outliers.

There are no planarity outliers.



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	R	194	0	98	6	0
2	Т	481	0	262	18	0
3	Ν	275	0	144	11	0
4	А	10828	0	10876	462	0
5	В	8874	0	8823	347	0
6	С	2101	0	2056	89	0
7	Ε	1731	0	1758	66	0
8	F	684	0	692	17	0
9	Н	1064	0	1029	56	0
10	Ι	952	0	897	41	0
11	J	532	0	542	30	0
12	Κ	919	0	929	49	0
13	L	337	0	352	14	0
14	А	2	0	0	0	0
14	В	1	0	0	0	0
14	С	1	0	0	0	0
14	Ι	2	0	0	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0
15	А	1	0	0	0	0
All	All	28981	0	28458	1053	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 18.

The worst 5 of 1053 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{l} \text{Interatomic} \\ \text{distance} \ (\text{\AA}) \end{array}$	Clash overlap (Å)
12:K:30:ALA:HA	12:K:75:ILE:O	1.74	0.88
4:A:613:ILE:HG21	9:H:102:TYR:HB3	1.57	0.85
4:A:1224:LEU:HD21	4:A:1240:CYS:HB3	1.60	0.84
4:A:109:HIS:CE1	4:A:110:CYS:SG	2.71	0.83
4:A:672:ASP:H	4:A:736:ASN:HD21	1.27	0.81

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	Perce	ntiles
4	А	1370/1733~(79%)	1341 (98%)	29~(2%)	0	100	100
5	В	1108/1224~(90%)	1088~(98%)	20~(2%)	0	100	100
6	С	265/318~(83%)	263~(99%)	2(1%)	0	100	100
7	Ε	210/215~(98%)	207~(99%)	3 (1%)	0	100	100
8	F	84/155~(54%)	82 (98%)	2(2%)	0	100	100
9	Η	129/146~(88%)	123~(95%)	6~(5%)	0	100	100
10	Ι	116/122~(95%)	114 (98%)	2(2%)	0	100	100
11	J	63/70~(90%)	62~(98%)	1 (2%)	0	100	100
12	Κ	112/120~(93%)	109~(97%)	3(3%)	0	100	100
13	L	41/70~(59%)	40 (98%)	1 (2%)	0	100	100
All	All	3498/4173 (84%)	3429 (98%)	69 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent side chain 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		
4	А	1194/1520~(79%)	1161 (97%)	33~(3%)	43	72	
5	В	955/1061~(90%)	929~(97%)	26 (3%)	44	73	
6	С	235/274~(86%)	227~(97%)	8 (3%)	37	68	
7	Е	193/197~(98%)	188 (97%)	5(3%)	46	74	

Continued on next page...



Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
8	F	73/137~(53%)	70~(96%)	3(4%)	30 63
9	Н	116/128 (91%)	112 (97%)	4 (3%)	37 68
10	Ι	110/116~(95%)	98~(89%)	12 (11%)	6 29
11	J	60/65~(92%)	58~(97%)	2(3%)	38 68
12	Κ	99/102~(97%)	98~(99%)	1 (1%)	76 88
13	L	37/57~(65%)	36~(97%)	1 (3%)	44 73
All	All	3072/3657~(84%)	2977 (97%)	95 (3%)	40 70

Continued from previous page...

5 of 95 residues with a non-rotameric side chain are listed below:

Mol	Chain	Res	Type
5	В	1182	CYS
7	Е	112	TYR
6	С	75	MET
6	С	199	LYS
9	Н	77	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
7	Е	115	ASN
9	Н	131	ASN
5	В	842	ASN
5	В	881	ASN
5	В	1161	HIS

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	7/9~(77%)	1 (14%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	2	U

There are no RNA pucker outliers to report.



5.4 Non-standard residues in protein, DNA, RNA chains (i)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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).

Mal	Mol Type Chai		Dog	Link	Bo	ond leng	\mathbf{ths}	B	ond ang	les
	туре	Ullalli	nes		Counts	RMSZ	# Z >2	Counts	RMSZ	# Z > 2
2	WVQ	Т	19	2	19,24,25	3.47	6 (31%)	20,33,36	1.53	4 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	WVQ	Т	19	2	-	4/6/40/41	0/2/2/2

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	Т	19	WVQ	C5-N7	12.49	1.46	1.28
2	Т	19	WVQ	C4-N9	4.83	1.45	1.35
2	Т	19	WVQ	C2-N2	4.04	1.45	1.34
2	Т	19	WVQ	C6-N1	-3.08	1.32	1.38
2	Т	19	WVQ	O6-C6	-3.05	1.18	1.23

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^{o})$	$Ideal(^{o})$
2	Т	19	WVQ	N3-C2-N1	-4.14	119.71	126.43
2	Т	19	WVQ	C2'-C3'-C4'	2.43	107.83	102.76
2	Т	19	WVQ	N2-C2-N3	2.14	120.04	116.57
2	Т	19	WVQ	N2-C2-N1	2.02	120.25	117.06

There are no chirality outliers.

All (4) torsion outliers are listed below:



Mol	Chain	Res	Type	Atoms
2	Т	19	WVQ	O4'-C4'-C5'-O5'
2	Т	19	WVQ	C3'-C4'-C5'-O5'
2	Т	19	WVQ	O4'-C1'-N9-C4
2	Т	19	WVQ	C4'-C5'-O5'-P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 9 ligands modelled in this entry, 9 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 (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	$\langle RSRZ \rangle$	#RSRZ>2	$\mathbf{OWAB}(\mathrm{\AA}^2)$	Q < 0.9
1	R	9/9~(100%)	-0.12	0 100 100	104, 120, 189, 222	0
2	Т	23/29~(79%)	-0.32	0 100 100	107, 185, 252, 281	0
3	N	13/18~(72%)	-0.34	0 100 100	203, 228, 274, 280	0
4	А	1384/1733~(79%)	0.01	44 (3%) 47 42	77, 117, 187, 248	0
5	В	1126/1224 (91%)	-0.06	16 (1%) 75 69	47, 108, 163, 210	0
6	С	267/318~(83%)	0.02	9 (3%) 45 40	46, 100, 155, 201	0
7	E	212/215~(98%)	0.10	19 (8%) 9 10	88, 137, 215, 274	0
8	F	86/155~(55%)	-0.25	1 (1%) 79 73	80, 108, 157, 183	0
9	Н	133/146~(91%)	0.37	11 (8%) 11 12	90, 137, 196, 280	0
10	Ι	118/122~(96%)	0.03	4 (3%) 45 40	93, 133, 178, 231	0
11	J	65/70~(92%)	-0.05	0 100 100	47, 100, 133, 150	0
12	K	114/120~(95%)	-0.15	2 (1%) 68 62	64, 108, 144, 179	0
13	L	43/70~(61%)	0.57	4 (9%) 8 9	94, 170, 235, 272	0
All	All	3593/4229~(84%)	-0.00	110 (3%) 49 43	46, 115, 186, 281	0

The worst 5 of 110 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
7	Ε	93	MET	7.8
7	Е	112	TYR	6.3
5	В	869	SER	5.6
4	А	1192	LEU	5.2
4	А	258	GLY	4.9



6.2 Non-standard residues in protein, DNA, RNA chains (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^{th} 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	$B-factors(Å^2)$	Q<0.9
2	WVQ	Т	19	23/24	0.84	0.27	140,145,169,176	0

6.3 Carbohydrates (i)

There are no monosaccharides 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^{th} 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	$B-factors(Å^2)$	Q < 0.9
14	ZN	А	1801	1/1	0.62	0.21	301,301,301,301	0
14	ZN	L	101	1/1	0.92	0.26	363,363,363,363	0
14	ZN	С	401	1/1	0.94	0.12	120,120,120,120	0
14	ZN	J	101	1/1	0.96	0.31	133,133,133,133	0
14	ZN	В	1301	1/1	0.96	0.07	204,204,204,204	0
14	ZN	Ι	202	1/1	0.97	0.15	273,273,273,273	0
14	ZN	Ι	201	1/1	0.98	0.10	113,113,113,113	0
14	ZN	А	1802	1/1	0.98	0.07	167,167,167,167	0
15	MG	А	1803	1/1	0.98	0.13	129,129,129,129	0

6.5 Other polymers (i)

There are no such residues in this entry.

