wwPDB X-ray Structure Validation Summary Report (i)

Dec 16, 2023 – 03:53 pm GMT

PDB ID	:	4AYB
Title	:	RNAP at 3.2Ang
Authors	:	Wojtas, M.N.; Mogni, M.; Millet, O.; Bell, S.D.; Abrescia, N.G.A.
Deposited on	:	2012-06-19
Resolution	:	3.20 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
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 (i)) were used in the production of this report:

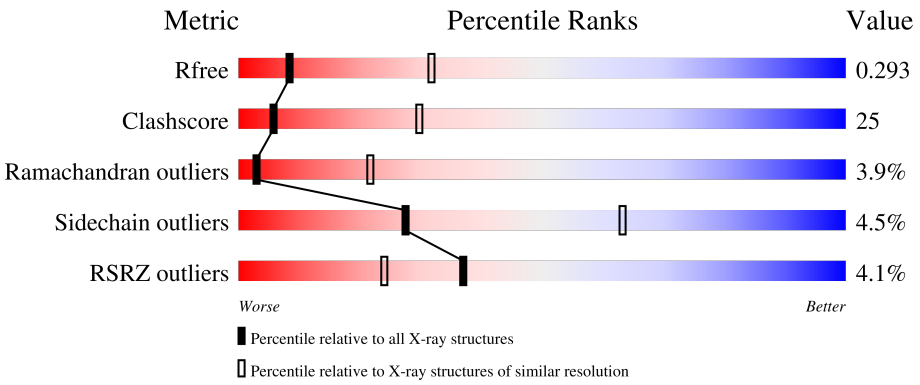
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.20 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	Quality of chain
1	A	880	2%, 55%, 39%, 5% ••
2	B	1131	2%, 59%, 34%, • •
3	C	395	6%, 47%, 42%, 6% • 5%
4	D	265	3%, 70%, 28%, ••
5	E	180	17%, 54%, 39%, •• 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	F	113	21% 40% 42% 11% 7%
7	G	132	3% 36% 43% 6% 14%
8	H	84	% 57% 31% 10%
9	K	95	% 63% 22% 12%
10	L	92	72% 27%
11	N	66	50% 41% 8%
12	P	48	4% 42% 42% 8% 8%
13	Q	104	8% 17% 25% 5% 52%

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
16	SF4	D	1264	-	-	X	-

2 Entry composition (i)

There are 16 unique types of molecules in this entry. The entry contains 54890 atoms, of which 27769 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.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	872	13986	4424	7029	1225	1282	26	0	0	0

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	1103	17664	5548	8908	1552	1627	29	0	0	0

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
3	C	376	5974	1840	3068	493	564	9	0	0	0

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
4	D	262	4215	1339	2128	337	398	13	0	0	0

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
5	E	171	2772	874	1413	229	251	5	0	0	0

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
6	F	105	1666	519	839	134	171	3	0	0	0

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
7	G	113	1816	572	915	152	173	4	0	0	0

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
8	H	76	1284	405	660	111	108	0	0	0

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
9	K	84	1390	431	717	123	118	1	0	0	0

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
10	L	91	1449	454	742	114	137	2	0	0	0

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
11	N	65	1058	332	537	94	88	7	0	0	0

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
12	P	44	744	236	387	62	54	5	0	0	0

- Molecule 13 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
13	Q	50	854	270	426	74	83	1	0	0	0

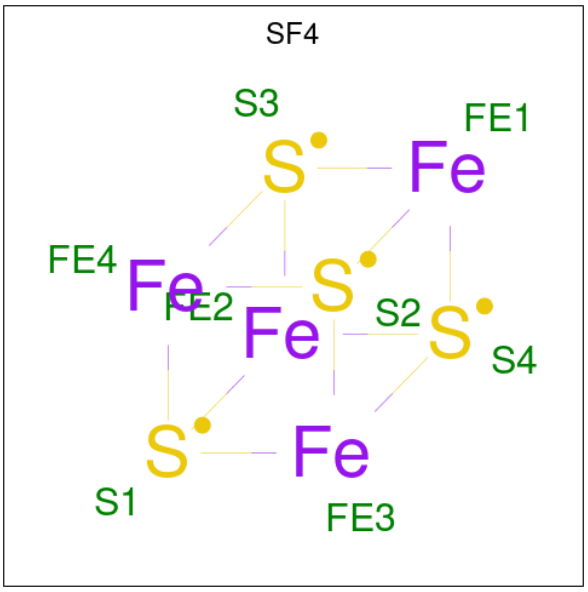
- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
14	A	4	4	4	0	0
			Total	Zn		
14	B	3	3	3	0	0
			Total	Zn		
14	C	1	1	1	0	0
			Total	Zn		
14	N	1	1	1	0	0
			Total	Zn		
14	P	1	1	1	0	0

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

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

- Molecule 16 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).



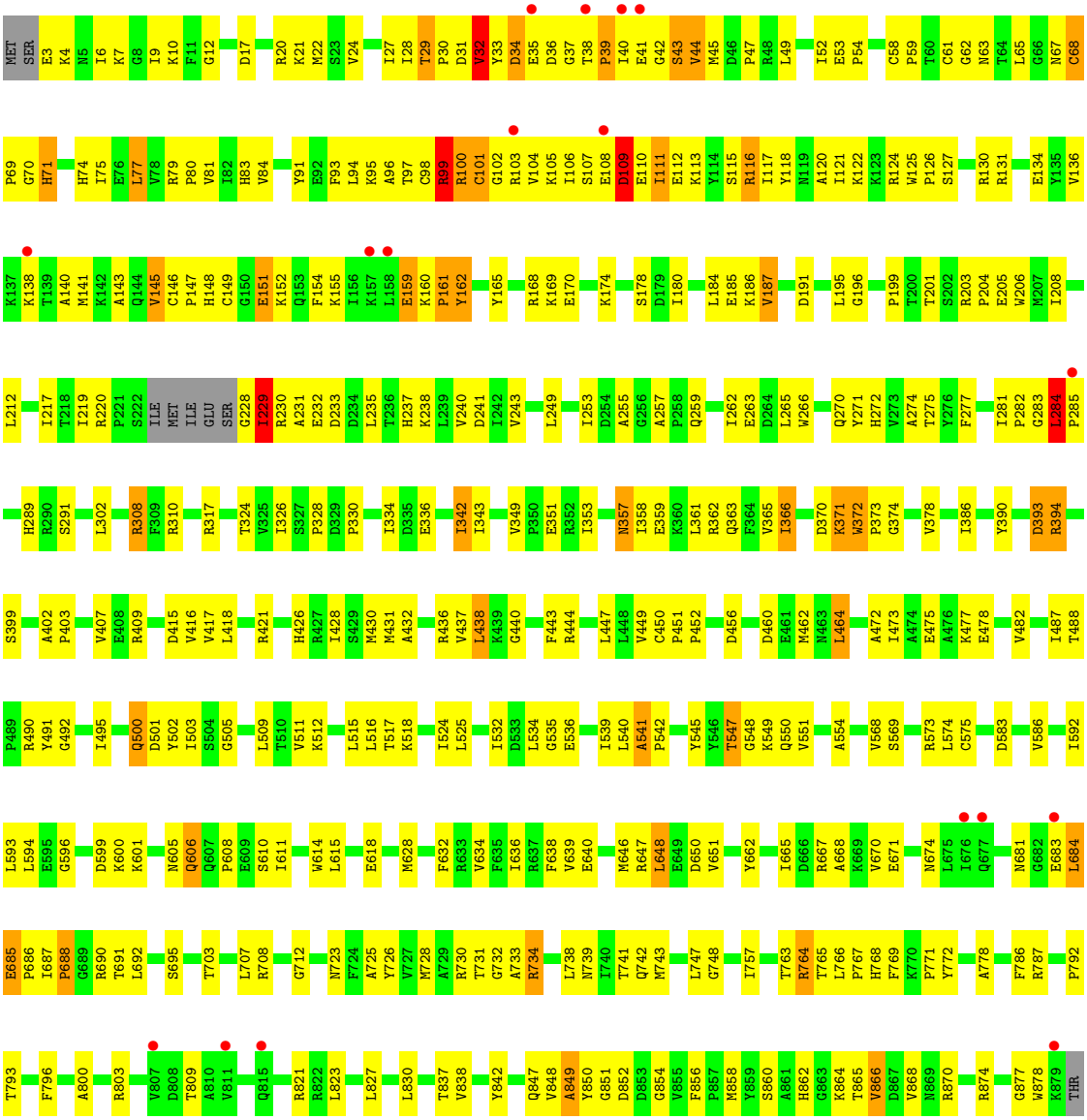
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	Fe	S		
16	D	1	7	3	4	0	0

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: DNA-DIRECTED RNA POLYMERASE

Chain A:



- Molecule 2: DNA-DIRECTED RNA POLYMERASE

Chain B: 2% 59% 34%

MET ASN GLU LEU SER SER N7 L8 S9 I10 D11 E12 R13 I17 Y20 F21 K24 G25 L26 V27 R28 L31 N35 V38 R39 N40 K41 L42 I52 P53 T54 E55 I56 L59 R62 L63 I66 R67 I68 R72 V73 R74 E75 E82 A88 R89 L90

R91 N92 L93 T94 Y95 L101 T102 M103 I104 N108 N109 I110 E111 A112 E113 P114 D121 M125 L126 D131 P132 I133 Y136 T137 L138 L141 D147 P148 K149 D150 Y154 F155 I156 E161 R162 T166 V175 T187 H188 T189 A190 K191 I192 I193 S194 S195 T196

A197 G198 Y199 R200 V201 P202 V203 R207 L208 K209 D210 G211 T212 F213 H214 F217 P218 A219 V220 P221 G222 K223 I224 P225 F226 V227 I228 L229 M230 L233 G234 I235 D238 R239 D240 I241 V242 Y243 A244 V245 S246 L247 I251 Q252 L255 F256 P257 S258 L259 A262 S263 N267

A271 L272 D273 S277 R278 V279 A280 I281 G282 Q283 K284 R285 E286 N287 R288 I289 E290 K291 Q294 K298 Y299 F300 L301 P302 T306 R311 A315 L318 A319 I322 S323 K324 V325 I326 L330 G331 R332 R333 E334 Y341 L348 L352 F353 L356 V359

A360 F361 F364 L368 L372 E373 K374 S375 K376 V377 R378 G379 R380 K381 L387 V388 H399 A400 L401 A402 T403 G404 N405 W406 V407 R410 Q415 R419 T420 N421 W422 L426 R431 R438 G439 Q440 P441 N442 A445 L448 Q452 W453 G454 C457

E460 T461 P462 N466 S467 G468 M476 A477 Q478 I479 A480 I483 N484 E485 K486 V500 V503 I504 G510 GLY GLU ASP GLN ASN E516 Y517 L518 K519 W520 S521 I524 L529 E537 E538 L539 I543 R544 E545 R546 E555 V558 G559 H560 T563 D564 F565 I566

N567 N572 R577 V578 R579 R580 P581 L582 I583 I584 V585 S586 N587 L591 V592 D596 I597 E598 K599 L600 S604 I605 T606 F607 D608 D609 L610 K615 I616 L619 D620 E624 Y628 V629 A630 L631 E632 P633 L636 H640 T641 H642 L643 W646 S647 P648 A649 I650

L651 G652 I653 T654 I658 P659 Y660 P661 Q665 S666 P667 Y671 Q672 S673 A674 M675 L680 G681 L682 Y683 A684 A685 N686 Y687 Q688 T691 R694 A695 H696 L697 L698 L705 V706 R709 A710 L711 P720 I726 I730 S731 F732 T733 G734 Y735 N736 M737 E738 D739 S740

I741 I742 M743 F757 F758 K767 Y768 P769 G770 G771 Q772 E773 D774 K775 I776 V777 M778 P779 E780 P781 G782 V783 R784 G785 Y786 K787 G788 K789 E790 Y791 Y792 K808 G809 R821 PHE LEU GLN GLU PHE LYS GLU LEU SER PRO GLU Q833 A834 K835 R836 I840 V841 T842 R843 H844

L852 V853 L854 I855 T856 E857 N862 R868 L872 R873 I874 P875 S876 I877 I891 D899 M900 P901 Y902 T903 V904 K905 G906 V907 V908 V911 I912 L913 H916 A917 L918 P919 M922 I927 M928 E929 G930 I931 K934 Y935 L938 I942 V943 F948 Y949 K950

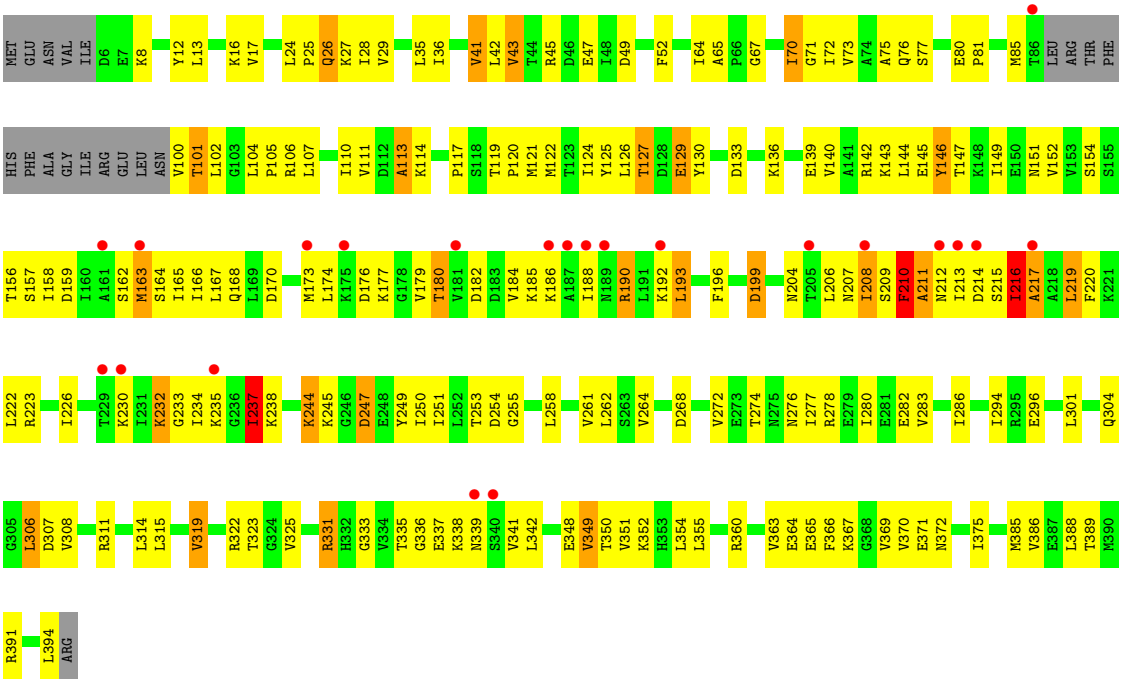
I963 E954 Q955 L956 Q957 I960 L961 Y965 L966 P967 D968 A969 T970 E971 V972 T973 Y974 D975 G976 R977 K981 I986 K995 L996 H1005 R1009 I1014 L1015 T1016 R1017 T1020 R1023 A1024 R1025 E1026 G1027 R1030 E1033 R1036 I1040 T1044 A1045 M1046 L1047

R1051 L1052 L1053 S1056 T1059 T1060 I1061 C1064 D1065 Q1066 C1067 I1070 G1071 W1072 K1075 N1076 K1077 N1078 K1079 Y1080 V1081 C1082 P1083 I1084 H1085 G1086 D1087 L1091 F1092 P1093 V1094 T1095 V1096 S1097 Y1098 K1101 L1102 L1103 I1104 Q1105 E1106 L1107 M1108 S1109 M1110 I1111 I1112 S1113 P1114 R1115 L1118 E1119

S1125 GLY GLY LYS GLY ASN GLU

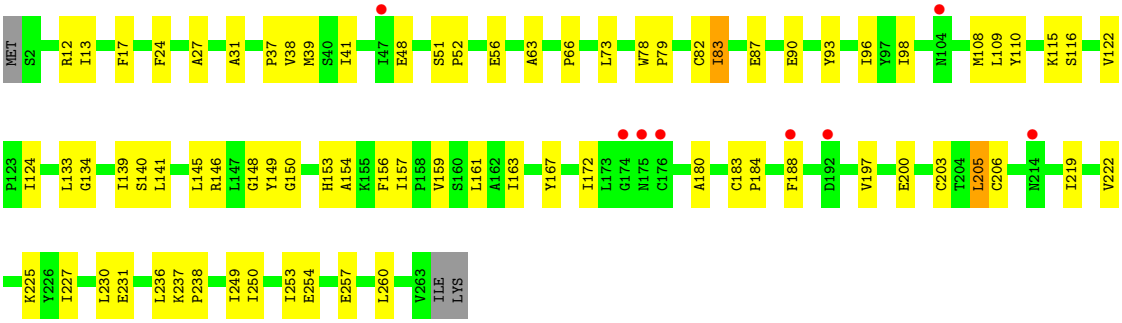
- Molecule 3: DNA-DIRECTED RNA POLYMERASE

Chain C: 6% 47% 42% 6% 5%



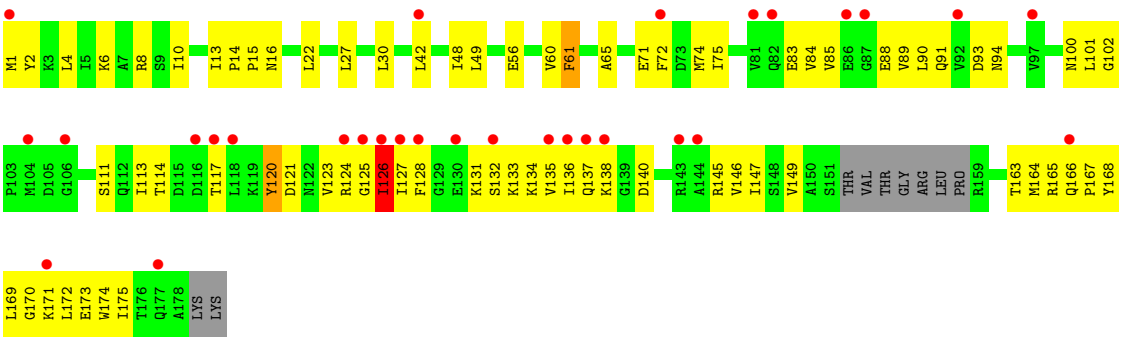
- Molecule 4: DNA-DIRECTED RNA POLYMERASE

Chain D:



- Molecule 5: DNA-DIRECTED RNA POLYMERASE

Chain E:



- Molecule 6: DNA-DIRECTED RNA POLYMERASE

Chain F: 21% 40% 42% 11% 7%

MET SER S3 V4 Y5 I6 V7 Y11 I12 P13 Y14 S15 V16 A17 L21 S27 G28 S29 S30 S31 L34 Q35 R36 V44 C47 S51 A52 Q53 K54 V55 E58 L59 S60 S61 I62 I63 S64 R65 V68 R69 A70 V71 L72 A73 S74 I75 C76 P77 I78 T79 P80

V83 R84 S85 I86 L87 I88 M89 D90 S91 N92 R93 T94 Y95 T96 S97 E98 D99 I100 Q101 K102 I103 I104 D105 I106 I107 ARG LYS TYR ILE LYS SER

- Molecule 7: DNA-DIRECTED RNA POLYMERASE

Chain G: 3% 36% 43% 6% 14%

MET MET GLU SER K5 A6 Q7 E8 I9 I10 L11 N16 S17 I18 E19 R20 L26 S27 I28 I29 C33 N34 D35 F36 S39 F40 D41 I42 I43 D44 S45 I46 N47 I48 Q51 V55 K56 A57 F58 I59 S60 K61 N62 R63 L64 S65 Y66 C72 I77 V78 T79 E80

S85 N88 Y92 I93 T94 I95 I96 S97 L98 F99 G100 L101 L102 V103 K104 I105 I106 S107 N108 K109 F112 L113 K114 I115 H116 Q117 LEU ASN VAL MET ASP HIS ILE TYR PHE CYS VAL LYS LYS ASN THR

- Molecule 8: DNA-DIRECTED RNA POLYMERASE

Chain H: % 57% 31% 10%

MET ARG GLY SER SER ASN ARG K8 R12 I13 H14 P18 K19 H20 L23 S24 I25 A28 R38 P39 L42 P43 W44 I45 D49 P50 V51 I55 K58 P59 I62 I63 R64 I65 I66 R67 K68 S69 V75 R79 Y80 V81 I82 S83 GLY

- Molecule 9: DNA-DIRECTED RNA POLYMERASE

Chain K: % 63% 22% 12%

MET GLY LEU GLU ARG ASP GLY ILE LEU SER Q11 W23 R34 V35 A47 I51 D52 I53 N54 N55 I56 T59 D60 V61 I62 S63 I64 E67 E68 R71 G72 V73 L74 T77 I88 S91 L92 R93 K94 SER

- Molecule 10: DNA-DIRECTED RNA POLYMERASE

Chain L: 72% 27%

M1 E2 I3 K4 I5 L6 N11 Y12 L13 E14 L15 E16 H22 L32 S39 H47 P48 V55 K56 I57 L58 T59 S62 L69 I73 R77 A80 Y83 I84 I87 T91 LYS

- Molecule 11: DNA-DIRECTED RNA POLYMERASE

Chain N: 50% 41% 8%

M1 M2 I3 P4 I5 R6 C7 F8 T9 C10 G11 I14 A15 W18 Q19 P20 L35 G39 V40 K41 R42 Y43 C44 C45 R46 R47 M48 L49 L50 S51 H52 I53 D54 I55 I56 I60 H61 R64 P65 ILE

- Molecule 12: DNA-DIRECTED RNA POLYMERASE

Chain P: 4% 42% 42% 8% 8%

MET ALA VAL TYR R5 C6 G7 K8 C9 W10 K11 T12 F13 T14 D15 E16 Q17 L18 L21 P22 G23 V24 R25 C26 C29 G30 Y31 K32 I33 I34 F35 M36 V37 R38 T41 I42 K43 I44 A47 I48

- Molecule 13: DNA-DIRECTED RNA POLYMERASE

Chain Q: 8% 17% 25% 5% 52%

MET VAL SER GLY MET SER THR GLU GLU GLU LYS GLU GLY THR ASN ASP GLU GLU VAL SER GLU GLU ARG GLU VAL GLU GLU THR SER GLU GLU GLU F33 P34 K35 L36 S37 I38 Q39 D40 I41 E42 L43 L44 M45 K46 N47 T48 E49 I50 L54 L55 N56 G57 K58 I59 S60 E63

L67 F68 E69 D70 N71 Y72 K73 D74 Y75 E76 K77 R78 D79 S80 R81 R82 LYS ALA LYS LYS ALA ALA SER LYS LYS VAL LYS LYS THR LYS LYS LYS GLU LYS SER VAL GLU GLY

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α, β, γ	195.72Å 212.41Å 128.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.10 – 3.20 40.10 – 3.20	Depositor EDS
% Data completeness (in resolution range)	99.4 (40.10-3.20) 99.5 (40.10-3.20)	Depositor EDS
R_{merge}	0.18	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.41 (at 3.18Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.242 , 0.300 0.235 , 0.293	Depositor DCC
R_{free} test set	4428 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	70.6	Xtriage
Anisotropy	0.333	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 69.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	54890	wwPDB-VP
Average B, all atoms (Å^2)	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.00% of the height of the origin peak. No significant pseudotranslation is detected.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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: MG, SF4, ZN

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	0.27	0/7108	0.45	0/9618
2	B	0.25	0/8923	0.44	0/12071
3	C	0.24	0/2930	0.46	0/3944
4	D	0.23	0/2123	0.39	0/2870
5	E	0.23	0/1379	0.40	0/1861
6	F	0.23	0/836	0.44	0/1133
7	G	0.24	0/913	0.43	0/1224
8	H	0.23	0/638	0.42	0/864
9	K	0.25	0/682	0.45	0/921
10	L	0.24	0/717	0.40	0/968
11	N	0.25	0/532	0.45	0/718
12	P	0.30	0/365	0.47	0/489
13	Q	0.23	0/434	0.41	0/580
All	All	0.25	0/27580	0.44	0/37261

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (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	6957	7029	7013	415	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	8756	8908	8888	418	0
3	C	2906	3068	3063	201	0
4	D	2087	2128	2125	59	0
5	E	1359	1413	1411	75	0
6	F	827	839	839	83	0
7	G	901	915	912	87	0
8	H	624	660	658	24	0
9	K	673	717	716	20	0
10	L	707	742	739	17	0
11	N	521	537	535	38	0
12	P	357	387	387	21	0
13	Q	428	426	426	38	0
14	A	4	0	0	0	0
14	B	3	0	0	0	0
14	C	1	0	0	0	0
14	N	1	0	0	0	0
14	P	1	0	0	0	0
15	A	1	0	0	0	0
16	D	7	0	0	4	0
All	All	27121	27769	27712	1393	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 25.

The worst 5 of 1393 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)
6:F:93:ARG:CB	6:F:94:THR:HA	1.87	1.05
1:A:146:CYS:SG	1:A:154:PHE:CZ	2.55	1.00
2:B:221:PRO:HB2	2:B:222:GLY:HA2	1.41	1.00
6:F:93:ARG:HB3	6:F:94:THR:HA	1.47	0.95
1:A:683:GLU:HA	1:A:684:LEU:HB3	1.51	0.93

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	868/880 (99%)	712 (82%)	124 (14%)	32 (4%)	3	22
2	B	1097/1131 (97%)	928 (85%)	133 (12%)	36 (3%)	4	25
3	C	372/395 (94%)	304 (82%)	49 (13%)	19 (5%)	2	15
4	D	260/265 (98%)	218 (84%)	37 (14%)	5 (2%)	8	39
5	E	167/180 (93%)	150 (90%)	16 (10%)	1 (1%)	25	64
6	F	103/113 (91%)	78 (76%)	17 (16%)	8 (8%)	1	6
7	G	111/132 (84%)	83 (75%)	20 (18%)	8 (7%)	1	7
8	H	74/84 (88%)	63 (85%)	8 (11%)	3 (4%)	3	21
9	K	82/95 (86%)	69 (84%)	9 (11%)	4 (5%)	2	17
10	L	89/92 (97%)	83 (93%)	4 (4%)	2 (2%)	6	35
11	N	63/66 (96%)	47 (75%)	10 (16%)	6 (10%)	0	3
12	P	42/48 (88%)	29 (69%)	9 (21%)	4 (10%)	0	3
13	Q	48/104 (46%)	38 (79%)	5 (10%)	5 (10%)	0	3
All	All	3376/3585 (94%)	2802 (83%)	441 (13%)	133 (4%)	3	22

5 of 133 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	34	ASP
1	A	43	SER
1	A	44	VAL
1	A	229	ILE
1	A	291	SER

5.3.2 Protein sidechains (i)

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	758/766 (99%)	717 (95%)	41 (5%)	22	58
2	B	951/975 (98%)	905 (95%)	46 (5%)	25	61
3	C	324/341 (95%)	307 (95%)	17 (5%)	23	59
4	D	235/238 (99%)	235 (100%)	0	100	100
5	E	150/158 (95%)	143 (95%)	7 (5%)	26	62
6	F	99/107 (92%)	90 (91%)	9 (9%)	9	34
7	G	106/125 (85%)	99 (93%)	7 (7%)	16	51
8	H	69/75 (92%)	68 (99%)	1 (1%)	67	86
9	K	74/83 (89%)	73 (99%)	1 (1%)	67	86
10	L	79/80 (99%)	79 (100%)	0	100	100
11	N	59/60 (98%)	57 (97%)	2 (3%)	37	70
12	P	40/43 (93%)	38 (95%)	2 (5%)	24	60
13	Q	48/96 (50%)	46 (96%)	2 (4%)	30	65
All	All	2992/3147 (95%)	2857 (96%)	135 (4%)	27	63

5 of 135 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
6	F	86	ILE
6	F	95	TYR
11	N	61	HIS
2	B	217	PHE
2	B	207	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	357	ASN
2	B	640	HIS
2	B	686	ASN
2	B	696	HIS
2	B	926	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 11 are monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
16	SF4	D	1264	4	0,9,12	-	-	-		

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
16	SF4	D	1264	4	-	-	0/3/3/5

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 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	D	1264	SF4	4	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	872/880 (99%)	0.05	17 (1%)	66	53	21, 75, 161, 226	0
2	B	1103/1131 (97%)	0.07	23 (2%)	63	49	26, 73, 154, 281	0
3	C	376/395 (95%)	0.38	22 (5%)	22	13	30, 89, 170, 288	0
4	D	262/265 (98%)	0.20	8 (3%)	49	32	54, 92, 150, 180	0
5	E	171/180 (95%)	1.03	30 (17%)	1	1	49, 169, 233, 275	0
6	F	105/113 (92%)	1.14	24 (22%)	0	0	103, 197, 244, 281	0
7	G	113/132 (85%)	0.56	4 (3%)	44	28	60, 123, 170, 266	0
8	H	76/84 (90%)	0.05	1 (1%)	77	65	49, 77, 132, 176	0
9	K	84/95 (88%)	-0.03	1 (1%)	79	67	35, 58, 149, 195	0
10	L	91/92 (98%)	0.05	0	100	100	48, 86, 132, 218	0
11	N	65/66 (98%)	0.08	0	100	100	36, 81, 119, 143	0
12	P	44/48 (91%)	0.27	2 (4%)	33	21	52, 94, 153, 227	0
13	Q	50/104 (48%)	0.90	8 (16%)	1	1	98, 154, 195, 204	0
All	All	3412/3585 (95%)	0.22	140 (4%)	37	24	21, 85, 188, 288	0

The worst 5 of 140 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	229	THR	7.3
6	F	78	ILE	6.4
1	A	40	ILE	5.9
6	F	79	THR	5.6
5	E	117	THR	5.4

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 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, 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	B-factors(Å^2)	Q<0.9
14	ZN	B	2128	1/1	0.83	0.29	81,81,81,81	0
14	ZN	A	1880	1/1	0.89	0.08	81,81,81,81	0
15	MG	A	1884	1/1	0.89	0.23	81,81,81,81	0
14	ZN	A	1882	1/1	0.92	0.34	81,81,81,81	0
14	ZN	A	1883	1/1	0.92	0.16	81,81,81,81	0
14	ZN	B	2127	1/1	0.93	0.34	81,81,81,81	0
14	ZN	C	1395	1/1	0.94	0.06	81,81,81,81	0
14	ZN	N	1066	1/1	0.94	0.21	88,88,88,88	0
14	ZN	A	1881	1/1	0.94	0.10	81,81,81,81	0
16	SF4	D	1264	7/8	0.95	0.21	81,81,81,81	0
14	ZN	B	2126	1/1	0.97	0.11	81,81,81,81	0
14	ZN	P	1049	1/1	0.98	0.10	118,118,118,118	0

6.5 Other polymers (i)

There are no such residues in this entry.