



# wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 28, 2017 – 02:22 PM EDT

PDB ID : 3H3V  
Title : Yeast RNAP II containing poly(A)-signal sequence in the active site  
Authors : Dengl, S.; Cramer, P.  
Deposited on : unknown  
Resolution : 4.00 Å(reported)

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

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20030345
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	rb-20030345

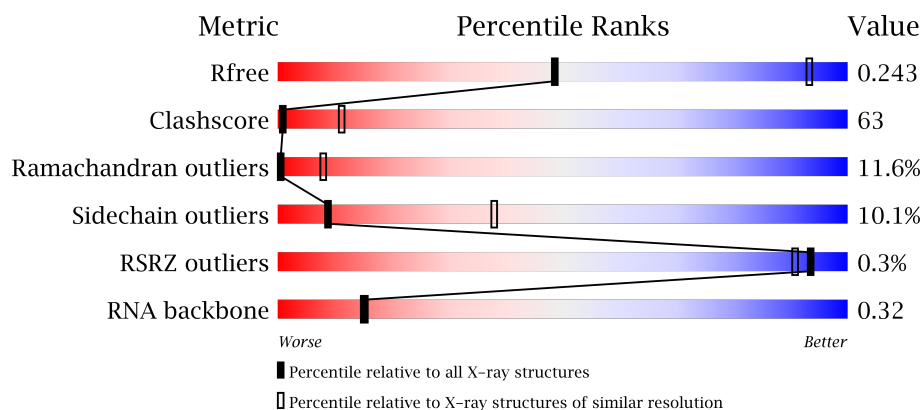
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


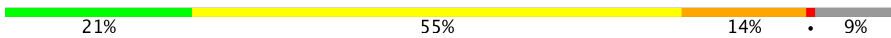


The reported resolution of this entry is 4.00 Å.

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.





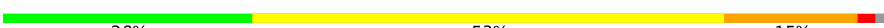


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	100719	1088 (4.40-3.60)
Clashscore	112137	1187 (4.40-3.60)
Ramachandran outliers	110173	1139 (4.40-3.60)
Sidechain outliers	110143	1126 (4.40-3.60)
RSRZ outliers	101464	1099 (4.40-3.60)
RNA backbone	2435	1026 (5.04-2.90)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	B	1733	
2	C	1224	
3	D	318	
4	E	221	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
5	F	215	
6	G	155	
7	H	171	
8	I	146	
9	J	122	
10	K	70	
11	L	120	
12	M	70	
13	N	14	
14	P	16	
15	T	26	

## 2 Entry composition [i](#)

There are 17 unique types of molecules in this entry. The entry contains 31777 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 II subunit RPB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	1416	Total	C	N	O	S	0	0	0
			11140	7021	1946	2111	62			

- Molecule 2 is a protein called DNA-directed RNA polymerase II subunit RPB2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1108	Total	C	N	O	S	0	0	0
			8810	5580	1541	1634	55			

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	266	Total	C	N	O	S	0	0	0
			2095	1317	348	417	13			

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	177	Total	C	N	O	S	0	0	0
			1427	882	256	287	2			

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	214	Total	C	N	O	S	0	0	0
			1752	1111	309	321	11			

- Molecule 6 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	84	Total	C	N	O	S	0	0	0
			679	434	115	127	3			

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	171	Total	C	N	O	S	0	0	0
			1340	861	222	249	8			

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	I	133	Total	C	N	O	S	0	0	0
			1068	673	180	211	4			

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	J	119	Total	C	N	O	S	0	0	0
			971	596	179	186	10			

- Molecule 10 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	K	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	L	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	M	46	Total	C	N	O	S	0	0	0
			364	224	72	64	4			

- Molecule 13 is a DNA chain called 5'-D(\*AP\*GP\*CP\*TP\*CP\*AP\*AP\*GP\*TP\*AP\*GP\*C  
P\*TP\*GP\*CP\*TP\*TP\*TP\*AP\*TP\*TP\*GP\*CP\*AP\*TP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	N	7	Total	C	N	O	P	11	0	0
			138	67	26	39	6			

- Molecule 14 is a RNA chain called 5'-D(\*CP\*AP\*GP\*CP\*TP\*AP\*CP\*TP\*TP\*GP\*AP\*GP\*CP\*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	P	8	Total	C	N	O	P	0	0	0
			168	77	33	51	7			

- Molecule 15 is a DNA chain called 5'-R(\*UP\*GP\*CP\*AP\*UP\*UP\*UP\*CP\*GP\*CP\*AP\*AP\*UP\*AP\*AP\*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	T	18	Total	C	N	O	P	8	0	0
			365	177	60	111	17			

- Molecule 16 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	J	2	Total	Zn	0	0
			2	2		
16	D	1	Total	Zn	0	0
			1	1		
16	K	1	Total	Zn	0	0
			1	1		
16	B	2	Total	Zn	0	0
			2	2		
16	C	1	Total	Zn	0	0
			1	1		
16	M	1	Total	Zn	0	0
			1	1		

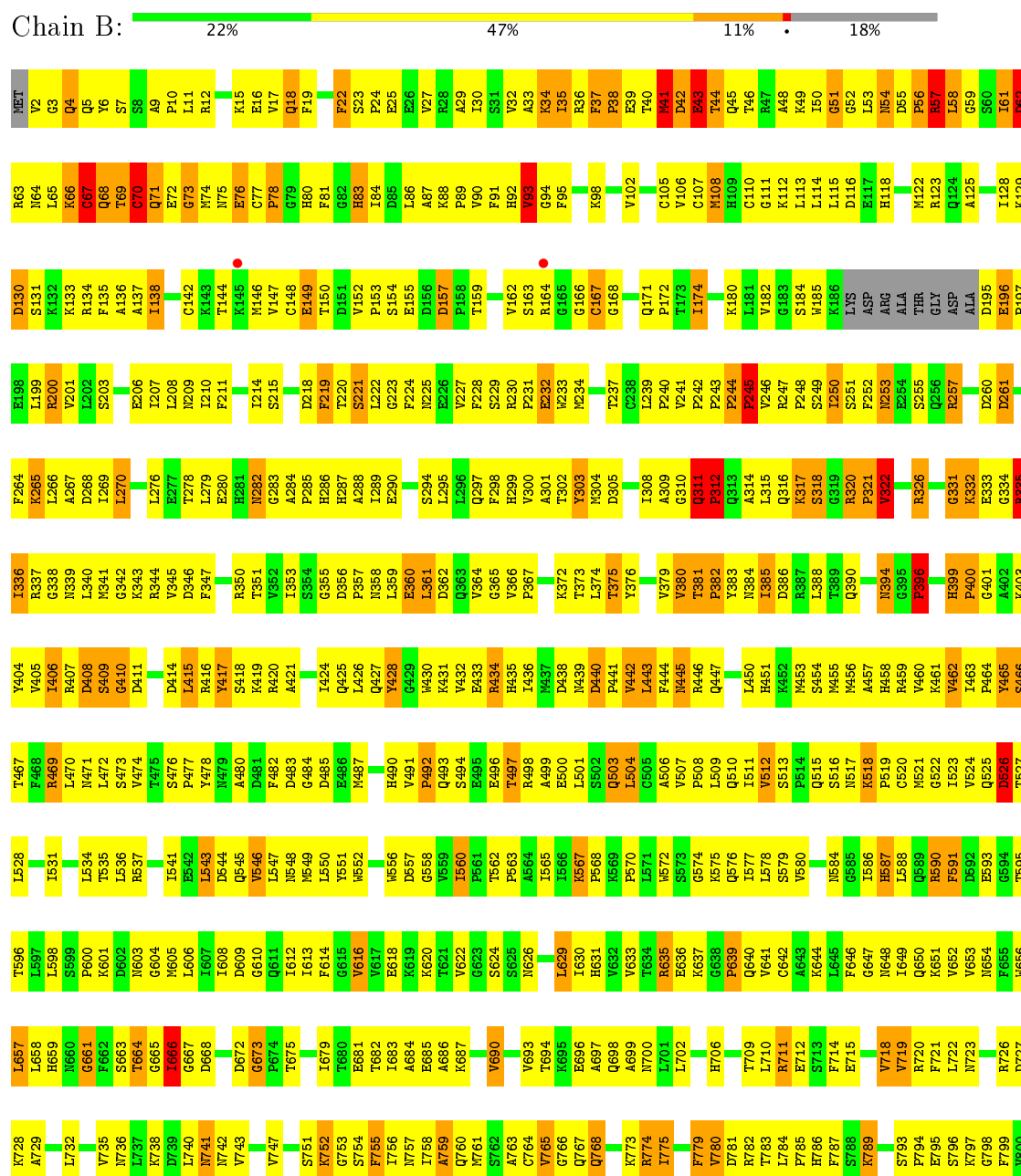
- Molecule 17 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	B	1	Total	Mg	0	0
			1	1		

### 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-directed RNA polymerase II subunit RPB1



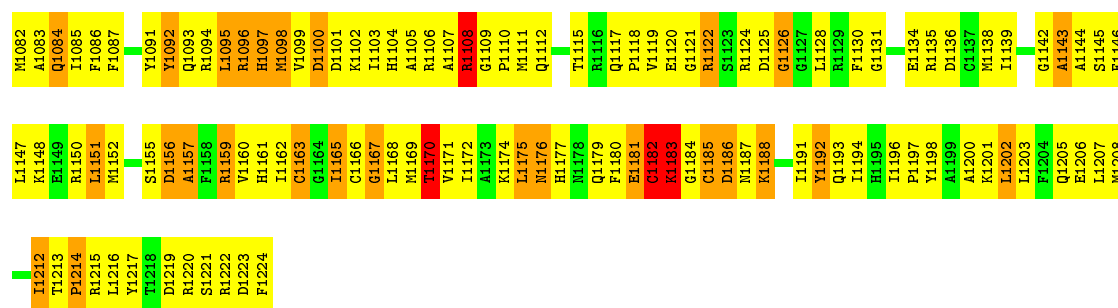
PRO	SER	SER	GLY	D1446	T1385	E1256	S1189	A1126	M1063	K1003	V937	I867	E801
SER	PRO	THR	PHE	E1447	R1386	M1259	L1192	D1127	V1064	M1004	R940	Y868	I802
THR	SER	ALA	THR	E1448	F1389	L1260	L1193	E1128	G1065	I1006	R940	G869	S803
ALA	THR	THR	THR	V1451	M1390	K1261	R1194	Q1130	V1066	I1007	K941	E870	T804
THR	PRO	THR	GLY		D1323	K1262	K1131	A1068	L1067	Q1008	L943	G872	R806
GLY	THR	GLY	GLY		D1324	L1263	L1197	K1132	L1072	M1009	R944	M873	G807
ALA	ALA	ALA	ALA		T1325	E1264	D1198	L1133	E1073	A1010	E945	D874	L808
ASP	ASP	ASP	ASP		R1326	E1265		L1134	A1074	Q1011	V946	A875	T809
GLY	GLY	GLY	GLY		I1327	M1266	A1201		P1075	R1012	F947	A876	
LYS	LYS	LYS	LYS		F1328	M1267	M1202		D1076	D1013	F948	H877	E812
THR	THR	THR	THR		M1329	L1268	A1137		A1076	A1014	I878	I878	F813
ALA	ALA	ALA	ALA		M1330	E1269	E1138		T1016	E1015			F814
THR	THR	THR	THR		S1331	M1270	E1139		Q1076	V1015			F815
SER	SER	SER	SER		F1332	L1271	H1140		L1017	A952			H816
PRO	PRO	PRO	PRO		I1333	T1272	T1141		M1079	N953			A817
THR	THR	THR	THR		G1401	T1273	T1142		L1080	M954			M818
GLY	GLY	GLY	GLY		F1402	L1273	L1143		C1019	P955			M819
ASP	ASP	ASP	ASP		R1274	K1274	K1144		ASN	L956			G819
ALA	ALA	ALA	ALA		M1336	G1275	S1145		THR	L1021			G820
GLY	GLY	GLY	GLY		E1337	V1276	V1146		PHE	L1022			R821
GLN	GLN	GLN	GLN		V1338	E1277	T1147		HIS	R1023			R821
ASP	ASP	ASP	ASP		L1339	M1278	E1214		PHE	S1024			R825
GLY	GLY	GLY	GLY		G1340	I1279	A1148		ALA	R1025			D826
VAL	VAL	VAL	VAL		L1409	E1280	E1151		GLY	L1026			T827
THR	THR	THR	THR		E1411	V1282	I1152		ALA	A1027			R836
PRO	PRO	PRO	PRO		E1412	M1284			SER	R1029			T829
TYR	TYR	TYR	TYR		G1413		Y1153		K1092	V1031			R830
ASN	ASN	ASN	ASN		A1414		D1155		L1093	L1032			T831
GLY	GLY	GLY	GLY		S1415		P1156		T1095	E1033			
GLU	GLU	GLU	GLU		E1416		D1157		S1096	E1034			T834
SER	SER	SER	SER		E1417		P1158		G1097	E1035			G835
VAL	VAL	VAL	VAL		L1418		A1159			R1036			
SER	SER	SER	SER		E1419		T1226		R1037	L1037			R839
GLY	GLY	GLY	GLY		D1420		I1227		H975	T907			R840
ASN	ASN	ASN	ASN		E1421		M1161		T976	L903			
ALA	ALA	ALA	ALA		R1422		V1163		K977	K1039			K843
ASP	ASP	ASP	ASP		G1423		P1164		P978	Q1040			L844
LEU	LEU	LEU	LEU		T1295		E1165		S979	L913			L845
ASP	ASP	ASP	ASP		E1296		D1166		D1042	E914			E846
VAL	VAL	VAL	VAL		E1297		E1167		L981	S915			D847
LYS	LYS	LYS	LYS		Y1298				T982	T848			T849
ASP	ASP	ASP	ASP		V1299				K983	M849			V850
THR	THR	THR	THR		K1300		I1170		K984	V851			H851
GLY	GLY	GLY	GLY		E1301		H1173		D855	L920			L852
LEU	LEU	LEU	LEU		P1302		F1174		I986	G921			D853
MET	MET	MET	MET		E1303		S1175		V987	D922			T854
PHE	PHE	PHE	PHE		Q1432		L1176		E1050	L923			T855
SER	SER	SER	SER		M1433		LEU		A1051	K924			T856
PRO	PRO	PRO	PRO		L1371		ARG		Q1052	L925			R857
ALA	ALA	ALA	ALA		V1372		PRO		F1053	Q926			H858
VAL	VAL	VAL	VAL		D1373		LYS		L1054	V927			S859
ASP	ASP	ASP	ASP		V1374		SER		R1055	L928			L860
THR	THR	THR	THR		M1375		LEU		S1056	N996			G861
GLY	GLY	GLY	GLY		T1376		ASP		V1057	L997			H862
SER	SER	SER	SER		G1377		ALA		H1059	E999			V863
ASN	ASN	ASN	ASN		Q1378		GLY		P1060	L1000			R864
ASP	ASP	ASP	ASP		F1441		THR		G1123	R934			Q865
ALA	ALA	ALA	ALA		D1442		GLU		H1124	G1002			F866
MET	MET	MET	MET		V1443		ASP		A1125				
VAL	VAL	VAL	VAL		M1444		Q1187						
ALA	ALA	ALA	ALA		I1445		Q1188						
GLY	GLY	GLY	GLY										

• Molecule 2: DNA-directed RNA polymerase II subunit RPB2

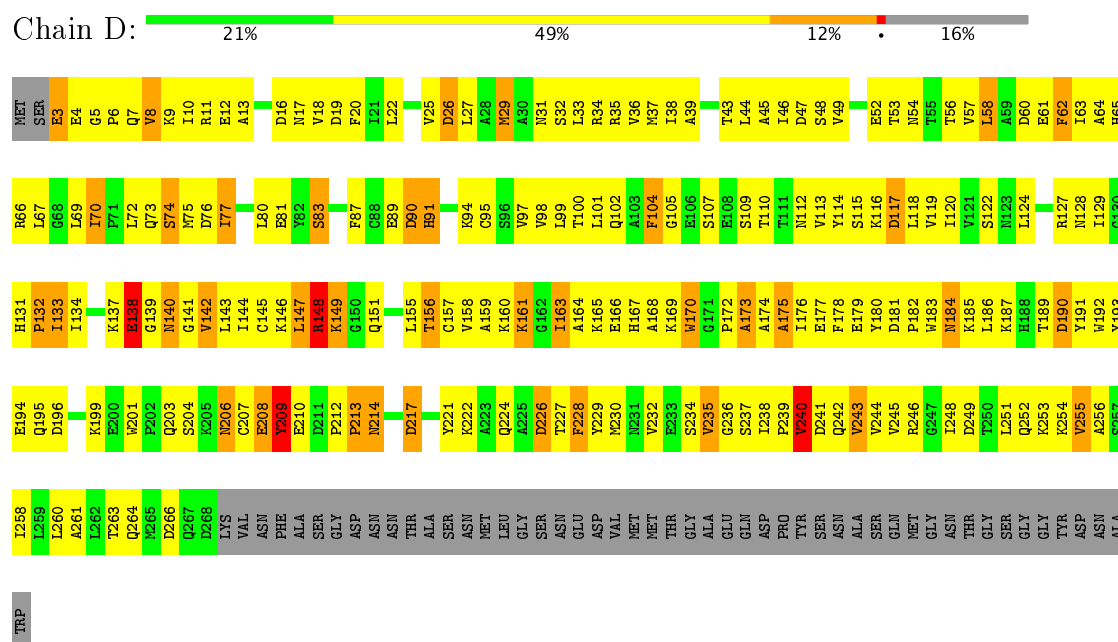
Chain C: 21% 55% 14% 9%



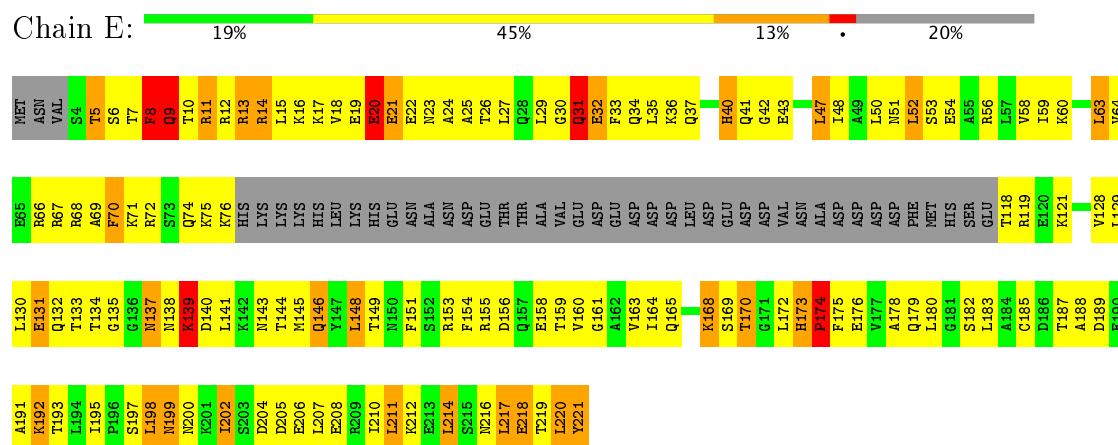
WORLDWIDE  
**PDB**  
PROTEIN DATA BANK



• Molecule 3: DNA-directed RNA polymerase II subunit RPB3

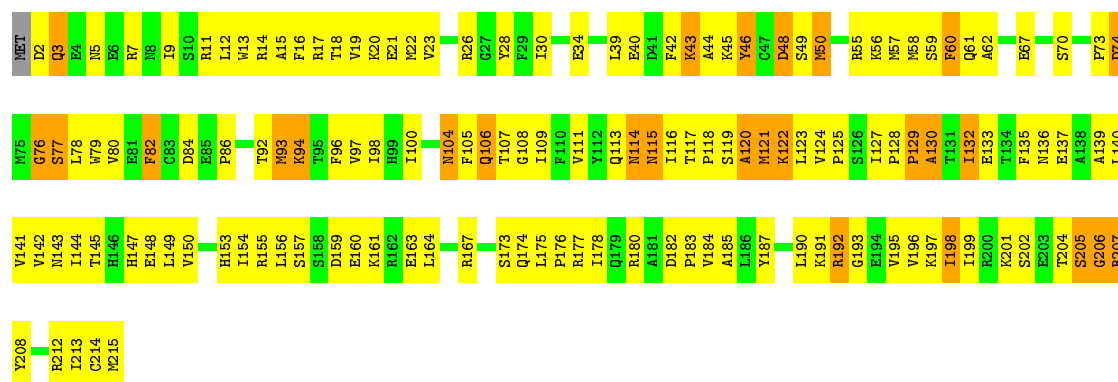


• Molecule 4: DNA-directed RNA polymerase II subunit RPB4

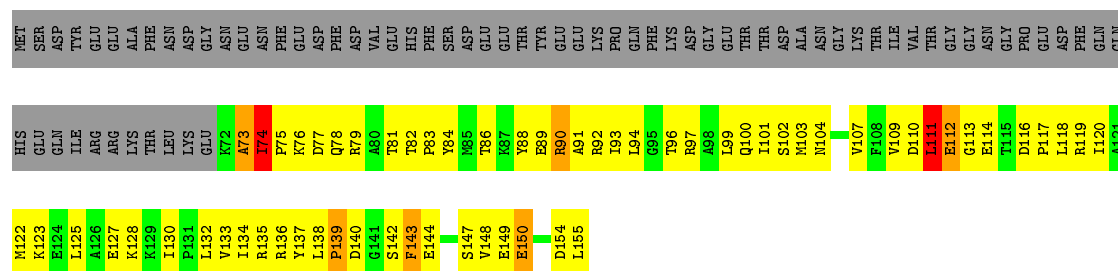


• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

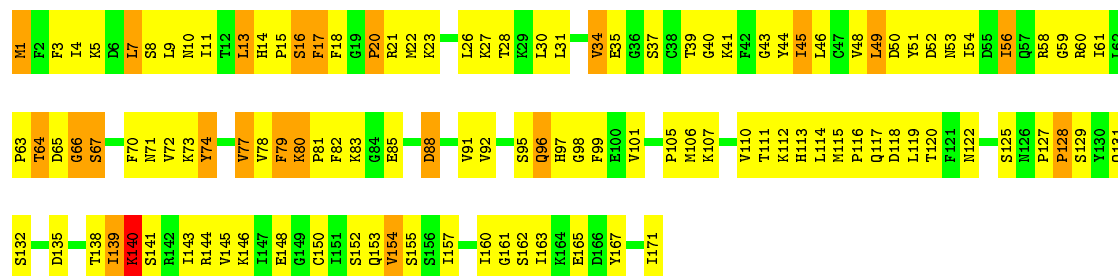




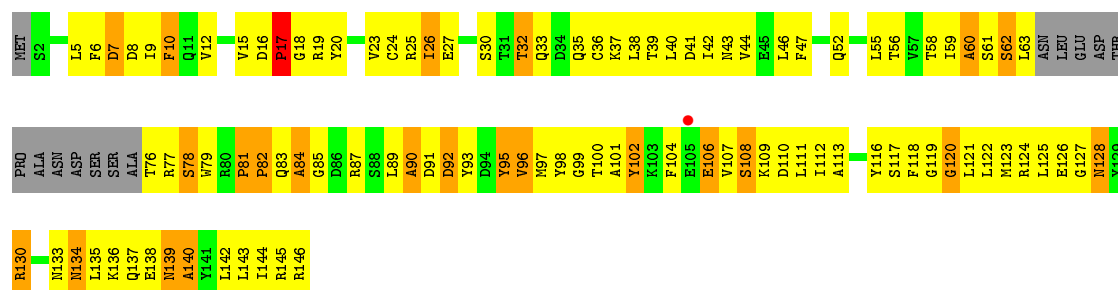
- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2



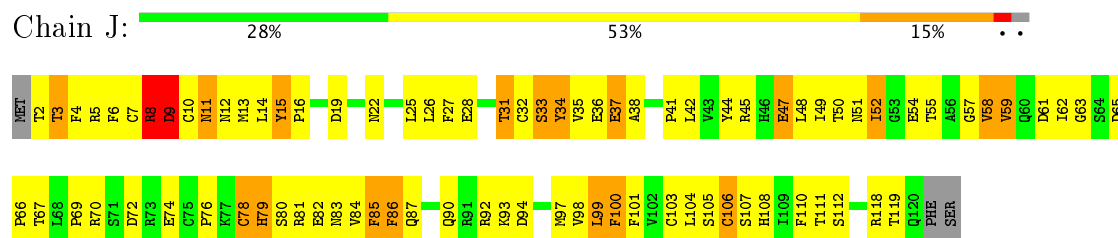
- Molecule 7: DNA-directed RNA polymerase II subunit RPB7



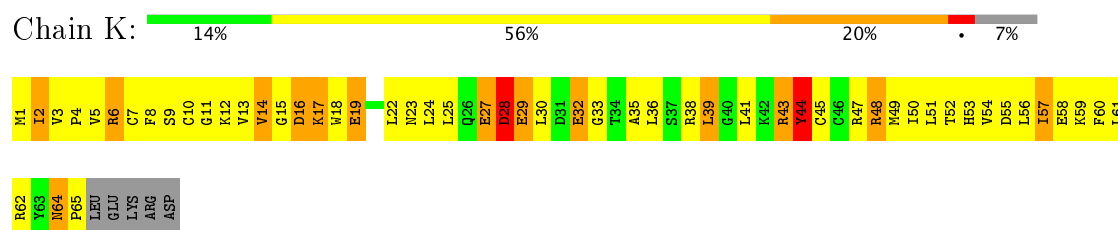
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3



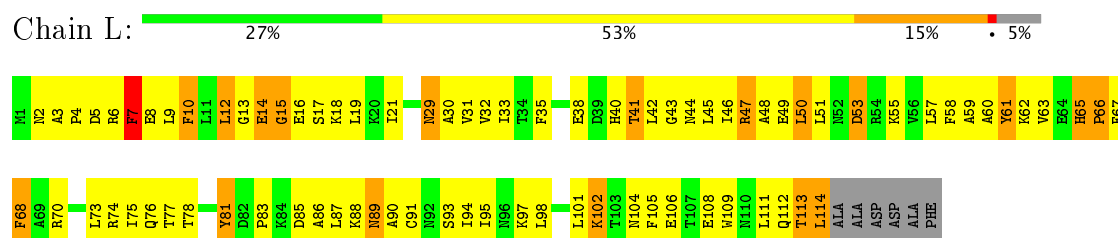
- Molecule 9: DNA-directed RNA polymerase II subunit RPB9



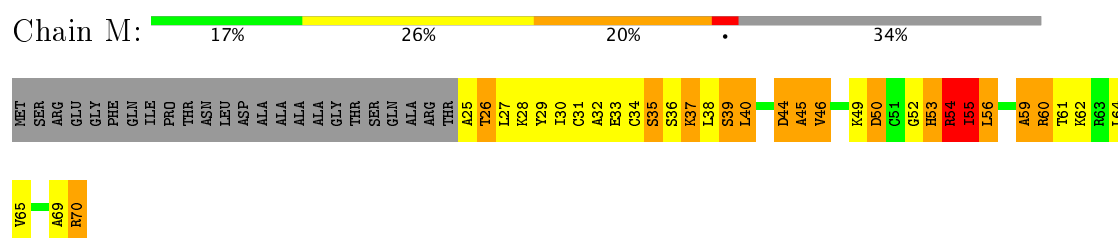
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5



- Molecule 11: DNA-directed RNA polymerase II subunit RPB11



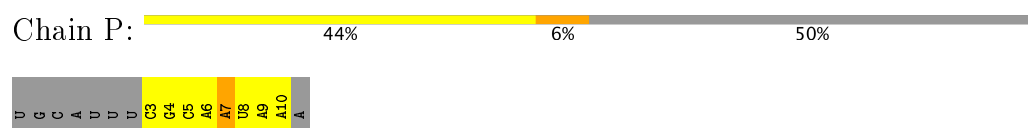
- Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4



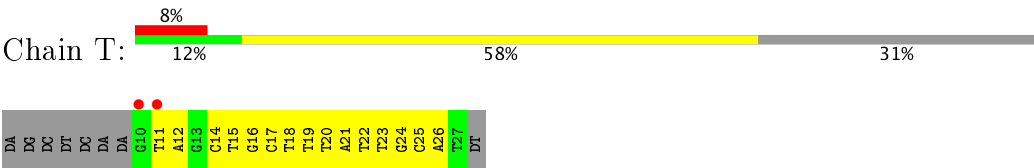
- Molecule 13: 5'-D(\*AP\*GP\*CP\*TP\*CP\*AP\*AP\*GP\*TP\*AP\*GP\*CP\*TP\*GP\*CP\*TP\*TP\*TP\*AP\*TP\*TP\*GP\*CP\*AP\*TP\*T)-3'



- Molecule 14: 5'-D(\*CP\*AP\*GP\*CP\*TP\*AP\*CP\*TP\*TP\*GP\*AP\*GP\*CP\*T)-3'



● Molecule 15: 5'-R(\*UP\*GP\*CP\*AP\*UP\*UP\*UP\*CP\*GP\*CP\*AP\*AP\*UP\*AP\*AP\*A)-3',



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	222.47Å 391.62Å 284.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 4.00 48.99 – 4.00	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-4.00) 100.0 (48.99-4.00)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.57 (at 4.00Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.205 , 0.241 0.215 , 0.243	Depositor DCC
$R_{free}$ test set	8343 reflections (7.99%)	DCC
Wilson B-factor (Å <sup>2</sup> )	131.8	Xtriage
Anisotropy	0.506	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.27 , 82.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.024 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.024 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	31777	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	138.0	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.*

<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 [i](#)

### 5.1 Standard geometry [i](#)

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	B	0.43	0/11339	0.71	4/15334 (0.0%)
2	C	0.43	0/8981	0.69	1/12108 (0.0%)
3	D	0.43	0/2133	0.71	0/2891
4	E	0.43	0/1437	0.69	0/1925
5	F	0.42	0/1788	0.67	0/2406
6	G	0.48	0/691	0.77	0/933
7	H	0.47	0/1368	0.73	0/1844
8	I	0.41	0/1086	0.69	0/1470
9	J	0.40	0/989	0.66	0/1331
10	K	0.44	0/541	0.75	0/727
11	L	0.45	0/937	0.69	0/1265
12	M	0.48	0/366	0.72	0/485
13	N	0.70	0/154	0.88	0/235
14	P	0.55	0/188	0.94	0/291
15	T	0.42	0/407	0.95	0/627
All	All	0.44	0/32405	0.71	5/43872 (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
10	K	0	1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	311	GLN	N-CA-C	5.60	126.11	111.00

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	56	PRO	N-CA-C	-5.54	97.71	112.10
1	B	425	GLN	N-CA-C	-5.38	96.48	111.00
2	C	1163	CYS	N-CA-C	-5.21	96.94	111.00
1	B	440	ASP	N-CA-C	-5.02	97.44	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
10	K	44	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	B	11140	0	11215	1481	0
2	C	8810	0	8847	1266	0
3	D	2095	0	2051	295	0
4	E	1427	0	1451	175	0
5	F	1752	0	1776	183	0
6	G	679	0	701	98	0
7	H	1340	0	1357	185	0
8	I	1068	0	1040	166	0
9	J	971	0	929	123	0
10	K	532	0	542	122	0
11	L	919	0	929	135	0
12	M	364	0	388	57	0
13	N	138	0	80	8	0
14	P	168	0	88	15	0
15	T	365	0	208	48	0
16	B	2	0	0	0	0
16	C	1	0	0	0	0
16	D	1	0	0	0	0
16	J	2	0	0	0	0
16	K	1	0	0	0	0
16	M	1	0	0	0	0
17	B	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	31777	0	31602	3977	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 63.

The worst 5 of 3977 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)
15:T:21:DA:C2'	15:T:22:DT:H5'	1.58	1.32
15:T:21:DA:H2''	15:T:22:DT:C5'	1.65	1.25
15:T:20:DT:C2'	15:T:21:DA:H5'	1.73	1.18
11:L:47:ARG:HH11	11:L:47:ARG:HB3	1.11	1.15
15:T:20:DT:H2'	15:T:21:DA:H5'	1.13	1.12

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	B	1406/1733 (81%)	965 (69%)	284 (20%)	157 (11%)	0	9
2	C	1090/1224 (89%)	719 (66%)	243 (22%)	128 (12%)	0	8
3	D	264/318 (83%)	163 (62%)	70 (26%)	31 (12%)	0	8
4	E	173/221 (78%)	107 (62%)	43 (25%)	23 (13%)	0	5
5	F	212/215 (99%)	154 (73%)	36 (17%)	22 (10%)	0	10
6	G	82/155 (53%)	62 (76%)	13 (16%)	7 (8%)	1	15
7	H	169/171 (99%)	129 (76%)	26 (15%)	14 (8%)	1	16
8	I	129/146 (88%)	79 (61%)	30 (23%)	20 (16%)	0	4
9	J	117/122 (96%)	80 (68%)	24 (20%)	13 (11%)	0	9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
10	K	63/70 (90%)	38 (60%)	12 (19%)	13 (21%)	0	2
11	L	112/120 (93%)	81 (72%)	24 (21%)	7 (6%)	1	24
12	M	44/70 (63%)	22 (50%)	8 (18%)	14 (32%)	0	0
All	All	3861/4565 (85%)	2599 (67%)	813 (21%)	449 (12%)	0	8

5 of 449 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	4	GLN
1	B	41	MET
1	B	48	ALA
1	B	54	ASN
1	B	57	ARG

### 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	B	1239/1520 (82%)	1113 (90%)	126 (10%)	8	37
2	C	962/1061 (91%)	867 (90%)	95 (10%)	9	39
3	D	234/274 (85%)	209 (89%)	25 (11%)	8	36
4	E	159/200 (80%)	136 (86%)	23 (14%)	4	24
5	F	196/197 (100%)	182 (93%)	14 (7%)	17	53
6	G	74/137 (54%)	69 (93%)	5 (7%)	18	55
7	H	152/152 (100%)	136 (90%)	16 (10%)	8	36
8	I	117/128 (91%)	107 (92%)	10 (8%)	12	47
9	J	113/116 (97%)	103 (91%)	10 (9%)	12	44
10	K	60/65 (92%)	54 (90%)	6 (10%)	9	38
11	L	99/102 (97%)	84 (85%)	15 (15%)	3	22
12	M	40/57 (70%)	36 (90%)	4 (10%)	9	38
All	All	3445/4009 (86%)	3096 (90%)	349 (10%)	9	38

5 of 349 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	502	ILE
2	C	999	MET
9	J	99	LEU
2	C	544	CYS
2	C	790	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 106 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	366	GLN
2	C	821	GLN
9	J	12	ASN
2	C	465	ASN
2	C	538	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
14	P	7/16 (43%)	1 (14%)	0

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
14	P	7	A

There are no RNA pucker outliers to report.

## 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 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

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

### 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	B	1416/1733 (81%)	-0.37	4 (0%) 93 91	76, 129, 175, 200	0
2	C	1108/1224 (90%)	-0.32	2 (0%) 94 93	79, 140, 185, 200	0
3	D	266/318 (83%)	-0.41	0 100 100	93, 126, 168, 188	0
4	E	177/221 (80%)	-0.34	0 100 100	106, 142, 182, 190	0
5	F	214/215 (99%)	-0.33	0 100 100	99, 161, 187, 200	0
6	G	84/155 (54%)	-0.52	0 100 100	73, 107, 138, 147	0
7	H	171/171 (100%)	-0.34	0 100 100	98, 127, 164, 174	0
8	I	133/146 (91%)	0.02	1 (0%) 86 79	134, 165, 186, 196	0
9	J	119/122 (97%)	-0.31	0 100 100	122, 165, 189, 200	0
10	K	65/70 (92%)	-0.59	0 100 100	92, 121, 158, 169	0
11	L	114/120 (95%)	-0.40	0 100 100	94, 128, 156, 170	0
12	M	46/70 (65%)	-0.03	0 100 100	120, 172, 195, 198	0
13	N	7/14 (50%)	1.19	2 (28%) 1 1	199, 200, 200, 200	1 (14%)
14	P	8/16 (50%)	0.05	0 100 100	198, 199, 200, 200	0
15	T	18/26 (69%)	0.65	2 (11%) 6 6	178, 199, 200, 200	1 (5%)
All	All	3946/4621 (85%)	-0.34	11 (0%) 93 91	73, 136, 184, 200	2 (0%)

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1176	LEU	3.7
15	T	11	DT	2.8
13	N	7	DC	2.8
2	C	471	LYS	2.8
8	I	105	GLU	2.3

## 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, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
16	ZN	J	123	1/1	0.99	0.16	0.27	136,136,136,136	0
16	ZN	K	71	1/1	0.99	0.25	0.14	113,113,113,113	0
16	ZN	D	319	1/1	0.99	0.13	-0.66	98,98,98,98	0
16	ZN	B	1735	1/1	0.99	0.12	-1.49	110,110,110,110	0
16	ZN	M	71	1/1	0.98	0.05	-1.61	174,174,174,174	0
16	ZN	B	1734	1/1	0.99	0.04	-2.66	149,149,149,149	0
16	ZN	J	124	1/1	0.92	0.03	-2.80	200,200,200,200	0
17	MG	B	1736	1/1	0.86	0.10	-	153,153,153,153	0
16	ZN	C	1225	1/1	0.99	0.23	-	104,104,104,104	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.