



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 1, 2023 – 03:31 PM EDT

PDB ID : 3HOX  
Title : Complete RNA polymerase II elongation complex V  
Authors : Sydow, J.F.; Brueckner, F.; Cheung, A.C.M.; Damsma, G.E.; Dengl, S.;  
Lehmann, E.; Vassylyev, D.; Cramer, P.  
Deposited on : 2009-06-03  
Resolution : 3.65 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.36  
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

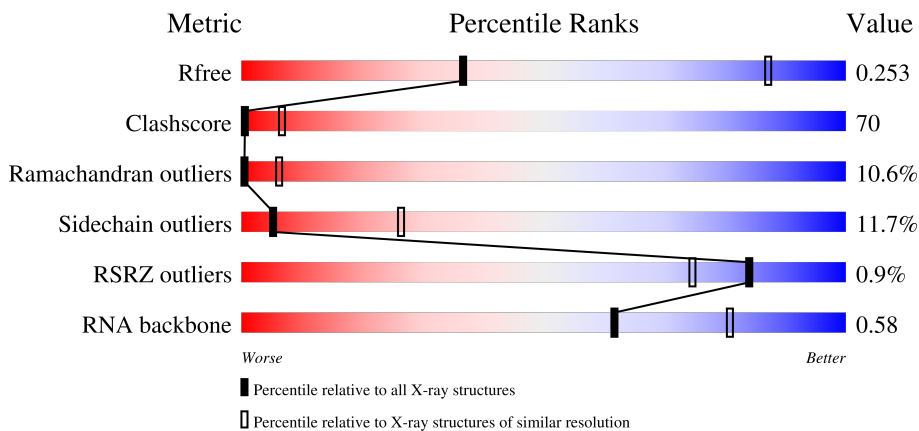
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.65 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.



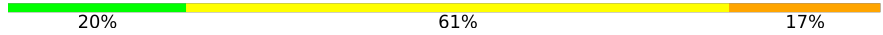
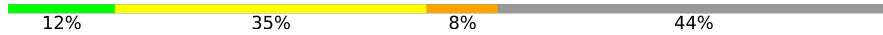

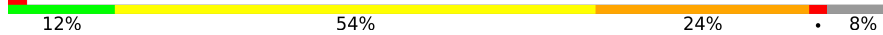
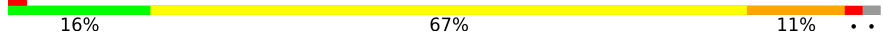
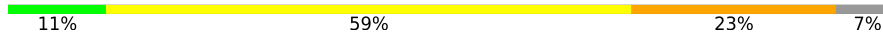
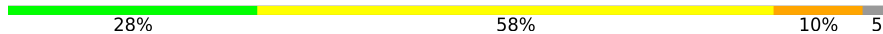

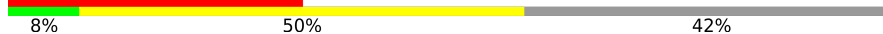
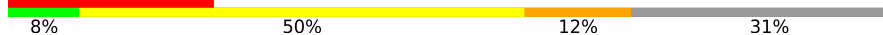

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1557 (3.82-3.50)
Clashscore	141614	1037 (3.80-3.52)
Ramachandran outliers	138981	1004 (3.80-3.52)
Sidechain outliers	138945	1002 (3.80-3.52)
RSRZ outliers	127900	1441 (3.82-3.50)
RNA backbone	3102	1024 (4.30-3.00)

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  $\geq 3$ , 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	1733	 21% 49% 11% 18%
2	B	1224	 20% 55% 15% 9%
3	C	347	 13% 48% 15% 23%
4	D	221	 19% 46% 15% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	E	215	
6	F	155	
7	G	171	
8	H	146	
9	I	122	
10	J	70	
11	K	120	
12	L	70	
13	N	12	
14	T	26	
15	P	18	

## 2 Entry composition i

There are 17 unique types of molecules in this entry. The entry contains 31918 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
			Total	C	N	O	S			
1	A	1417	11151	7027	1950	2112	62	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	1113	8847	5600	1552	1640	55	0	0	0

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

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

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-28	MET	-	expression tag	UNP P16370
C	-27	GLY	-	expression tag	UNP P16370
C	-26	SER	-	expression tag	UNP P16370
C	-25	HIS	-	expression tag	UNP P16370
C	-24	HIS	-	expression tag	UNP P16370
C	-23	HIS	-	expression tag	UNP P16370
C	-22	HIS	-	expression tag	UNP P16370
C	-21	HIS	-	expression tag	UNP P16370
C	-20	HIS	-	expression tag	UNP P16370
C	-19	SER	-	expression tag	UNP P16370
C	-18	ASN	-	expression tag	UNP P16370
C	-17	SER	-	expression tag	UNP P16370
C	-16	GLY	-	expression tag	UNP P16370
C	-15	LEU	-	expression tag	UNP P16370

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-14	ASN	-	expression tag	UNP P16370
C	-13	ASP	-	expression tag	UNP P16370
C	-12	ILE	-	expression tag	UNP P16370
C	-11	PHE	-	expression tag	UNP P16370
C	-10	GLU	-	expression tag	UNP P16370
C	-9	ALA	-	expression tag	UNP P16370
C	-8	GLN	-	expression tag	UNP P16370
C	-7	LYS	-	expression tag	UNP P16370
C	-6	ILE	-	expression tag	UNP P16370
C	-5	GLU	-	expression tag	UNP P16370
C	-4	TRP	-	expression tag	UNP P16370
C	-3	HIS	-	expression tag	UNP P16370
C	-2	GLU	-	expression tag	UNP P16370
C	-1	ASP	-	expression tag	UNP P16370
C	0	THR	-	expression tag	UNP P16370
C	1	GLY	-	expression tag	UNP P16370

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	178	1434	887	257	288	2	0	0	0

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
6	F	87	705	451	119	132	3	0	0	0

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
8	H	135	1081	680	183	214	4	0	0	0

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

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

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

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

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

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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
12	L	46	363	224	72	63	4	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
13	N	7	137	68	22	41	6	0	0	0

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	Br	C	N	O	P			
14	T	18	369	1	176	69	106	17	0	0	0

- Molecule 15 is a RNA chain called 5'-R(\*UP\*GP\*CP\*AP\*UP\*UP\*U\*CP\*AP\*AP\*CP\*CP

\*AP\*GP\*GP\*CP\*UP\*U)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
15	P	11	213	95	39	69	10	0	0	0

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

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
16	A	2	Total 2	Zn 2	0	0
16	B	1	Total 1	Zn 1	0	0
16	C	1	Total 1	Zn 1	0	0
16	I	2	Total 2	Zn 2	0	0
16	J	1	Total 1	Zn 1	0	0
16	L	1	Total 1	Zn 1	0	0

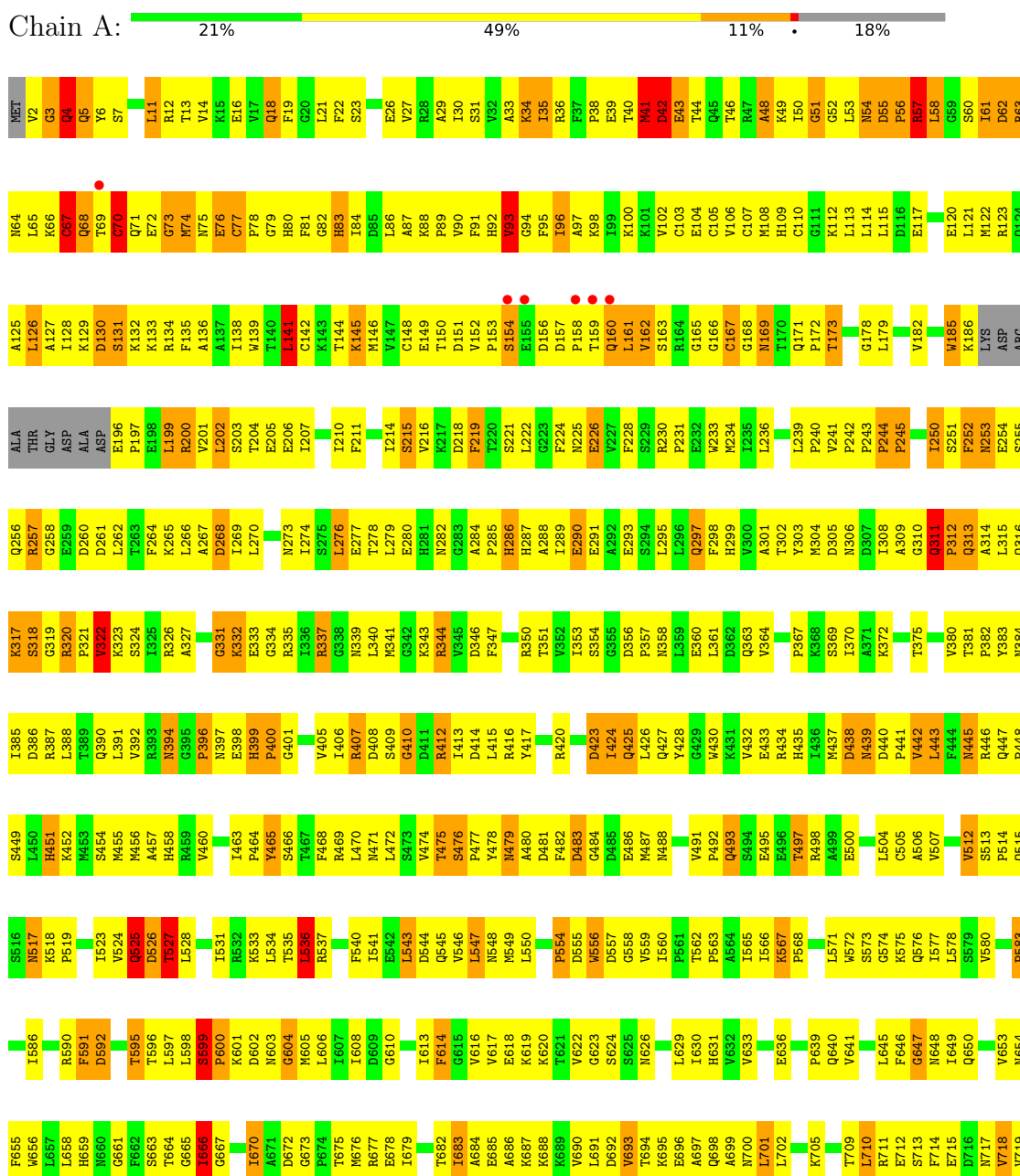
- Molecule 17 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

### 3 Residue-property plots

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 II subunit RPB1





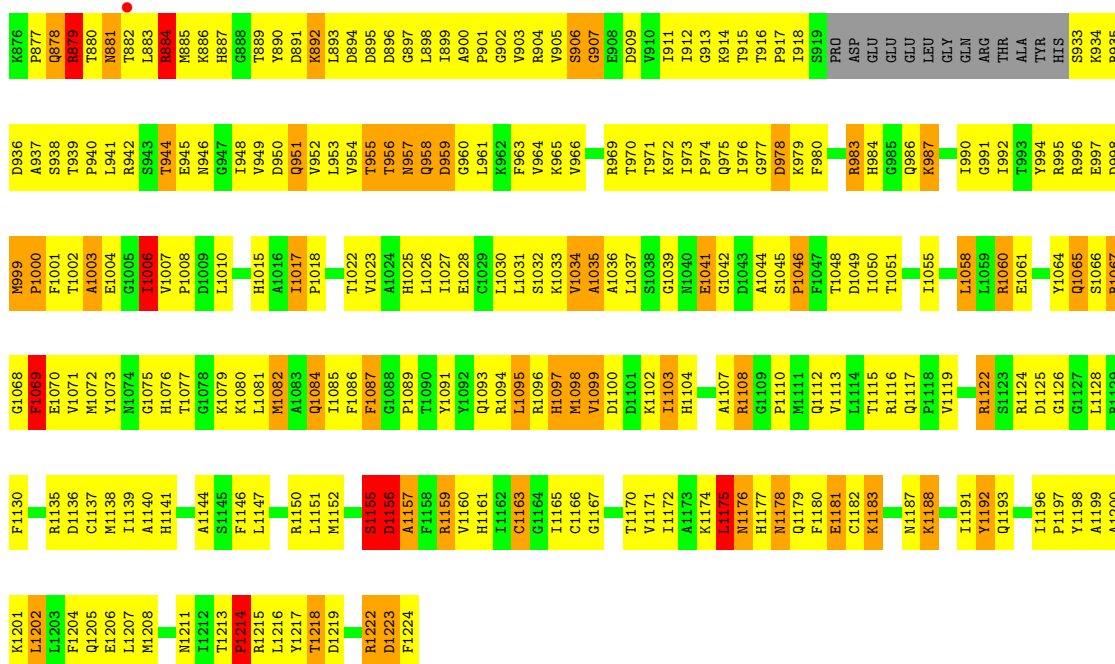


ARG

Molecule 2: DNA-directed RNA polymerase II subunit RPB2

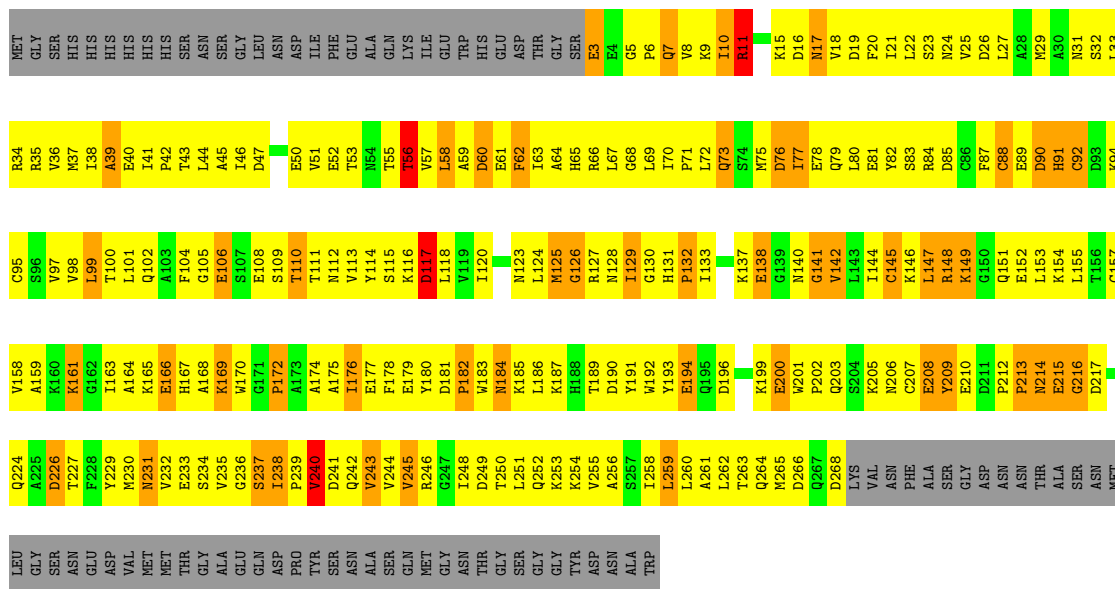


MET	SER	ASP	LEU	ALA	ASN	SER	GLU	LYS	TYR	TYR	ASP	GLU	ASP	PRO	TYR	GLY	PHE	GLU	D20	E21	S22	A23	P24	I25	T26	A27	E28	D29	S30	A31	A32	V33	I34	S35	A36	F37	F38	R39	E40	E42	L43	V44	S45	Q46	Q47	L48	D49	H50	F51	N52	O53	F54	V55	D56	Y57	L58	L59	Q60							
D61	I62	I63	C64	E65	D66	S67	I68	L69	I70	LEU	GLN	GLU	ASP	LEU	ALA	GLN	ALA	HIS	THR	PHE	THR	GLU	D20	E21	S22	A23	P24	I25	T26	A27	E28	D29	S30	A31	A32	V33	I34	S35	A36	F37	F38	R39	E40	E42	L43	V44	S45	Q46	Q47	L48	D49	H50	F51	N52	O53	F54	V55	D56	Y57	L58	L59	Q60			
T123	Y124	T185	E186	S126	G127	L128	F129	L130	D131	V132	K133	K134	ARG	THR	TYR	GLN	ALA	GLY	HIS	THR	PHE	THR	GLU	D20	E21	S22	A23	P24	I25	T26	A27	E28	D29	S30	A31	A32	V33	I34	S35	A36	F37	F38	R39	E40	E42	L43	V44	S45	Q46	Q47	L48	D49	H50	F51	N52	O53	F54	V55	D56	Y57	L58	L59	Q60		
T185	E186	S126	G127	L128	F129	L130	D131	V132	K133	K134	ARG	THR	TYR	GLN	ALA	GLY	HIS	THR	PHE	THR	GLU	D20	E21	S22	A23	P24	I25	T26	A27	E28	D29	S30	A31	A32	V33	I34	S35	A36	F37	F38	R39	E40	E42	L43	V44	S45	Q46	Q47	L48	D49	H50	F51	N52	O53	F54	V55	D56	Y57	L58	L59	Q60				
M199	G201	Y202	F203	I204	I205	G206	R261	E262	S265	K266	L267	L268	T272	L273	G274	S275	I276	K277	L278	D279	I280	L281	I282	V283	A284	G285	F286	R287	A288	L289	D290	S291	I292	P293	D294	G295	E296	L297	L298	H299	H300	Y303	D304	V305	N306	D307	F370	E371	S372	R373	Q374	L311													
E312	M313	L314	K315	L316	C317	Y318	V323	I324	K385	D326	R327	E328	T329	A330	L331	D332	F333	I334	G335	R336	ARG	THR	ALA	GLY	LEU	LYS	K345	A346	K347	V348	L349	A350	G351	L352	Q353	E354	F355	L356	K357	E358	F360	L361	P362	H363	L364	Y365	Q366	L367	E368	G369	F370	E371	S372	R373	Q374										
A375	F376	F377	G378	G379	Y380	M381	L382	N383	R384	Y385	L386	L387	C388	A389	L390	D391	R392	F393	D394	Q395	D396	R397	R398	Y399	H400	F401	G402	K403	G404	M405	L406	D407	L408	A409	G410	P411	L412	L413	A414	Q415	F416	K417	T418	L419	L420	F421	K422	K423	L424	T425	K426	D427	I428	F429	R430	Y431	M432	Q433	R434						
T435	V436	E437	GLU	ALA	HIS	ASP	PHE	ASN	MET	LYS	L446	A447	I448	M449	A450	K451	T452	T453	G454	L455	T456	L457	K458	Y459	L460	L461	A462	G463	G464	M465	M466	A467	E468	Q469	K470	A471	K472	A473	S474	S475	R476	A477	S480	Q481	V482	L483	M484	R485	Y486	T487	Y488	L489	T490	L491	L492	L493	L494	L495	R496	L497	T498				
M499	T500	P501	I502	D505	G506	K507	S508	A509	D510	K511	P512	H513	L514	H515	N516	K517	T518	H519	L520	M521	E522	R523	Y524	P525	E526	F527	G528	V529	M530	K531	M532	M533	L534	S535	S536	S537	M538	L539	S540	L541	M542	S543	G544	L545	S546	G547	T548	D549	P550	F551	M552	P553	L554	M555	T556	F557	L558	S559	E560	V561	G562	M563	E564	P565	L566
E567	D568	Y569	V570	P571	H572	O573	S574	P575	D576	A577	T578	R579	V580	F581	V582	N583	G584	V585	M586	H587	L588	E589	H590	R591	N592	F593	A594	R595	L596	M597	E598	T599	L600	R601	T602	L603	R604	R605	G606	D607	P608	M609	G610	P611	E612	V613	S614	L615	M616	R617	D618	L619	R620	E621	K622	E623	L624	G625	P626	G627					
F627	T628	D629	A630	G631	R632	V633	T634	P635	F636	L637	F638	L639	V640	E641	D642	D643	E644	S645	L646	H648	H649	E650	L651	K652	V653	R654	K655	A656	L657	L658	L661	M662	A663	T664	E665	T666	Q667	I668	GLY	PHE	GLY	GLY	PHE	GLU	GLU	GLU	VAL	GLU	E676	V679	T680	M681	R682	M683	L684	L685	M686	E687	G688						
L689	V690	E691	G692	D693	D694	A695	E696	Q697	R698	V699	S700	I703	A704	M705	Q706	P707	E708	D709	L710	E711	P712	A715	ASN	GLU	GLY	L781	L782	S783	A784	L785	L786	L787	R788	I789	R790	L791	S792	H793	L794	A795	L796	L797	F798	P799	Q800	R801	P802	L803	K804	G805	T806	R806	R807	A808	S809	I810	E811	L812	F813	L814	R815				
I755	I756	F757	F758	P759	D760	H761	M762	Q763	Q764	S765	P766	T767	T768	Y769	Q770	S771	A772	M773	G774	K775	E776	M777	A778	L779	L780	L781	L782	L783	L784	L785	L786	L787	L788	L789	L790	L791	L792	L793	L794	L795	L796	L797	L798	L799	L800	L801	L802	L803	L804	L805	L806	L807	L808	L809	L810	L811	L812	L813	L814	L815					
E816	L817	P818	A819	G820	Q821	N822	A823	I824	A825	P826	I827	R828	C829	Y830	S831	G832	Y833	N834	Q835	E836	D837	S838	M839	I840	M841	N842	Q843	S844	N845	L846	D847	R848	G849	L850	F851	R852	S853	L854	F855	F856	R857	S858	Y859	M860	D861	Q862	E863	K864	K865	Y866	G867	M868	S869	T870	R871	E872	L873	F874	E875						



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

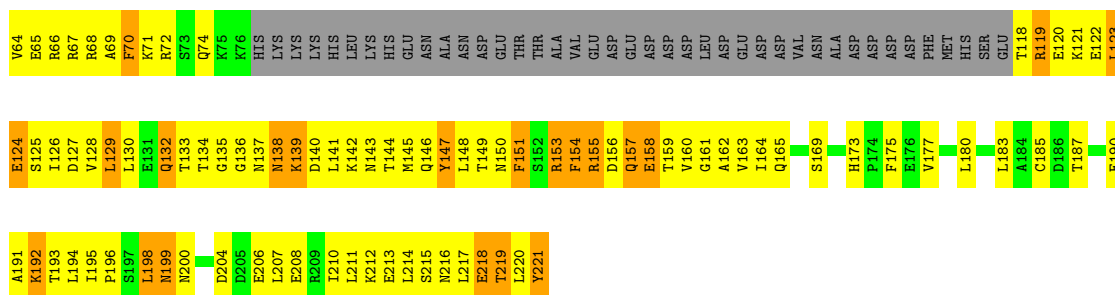
Chain C: 13% 48% 15% 23%



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

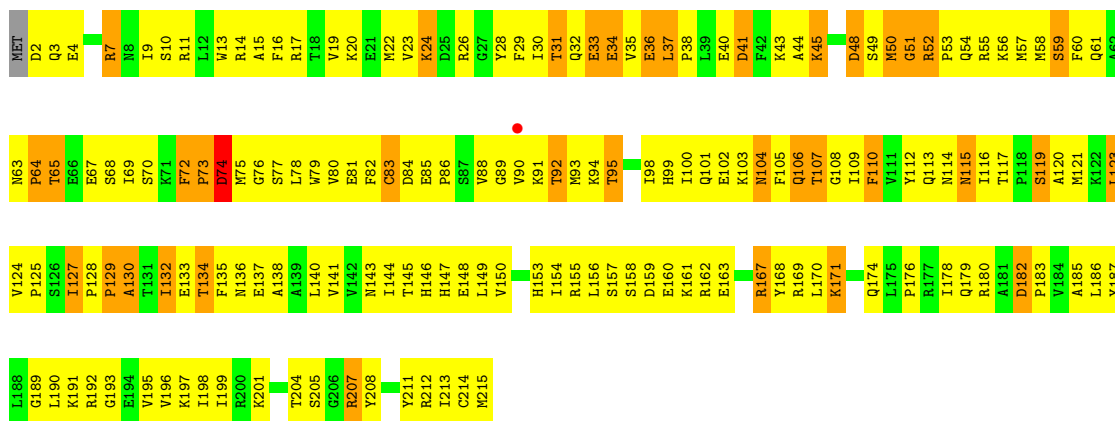
Chain D: 19% 46% 15% 19%





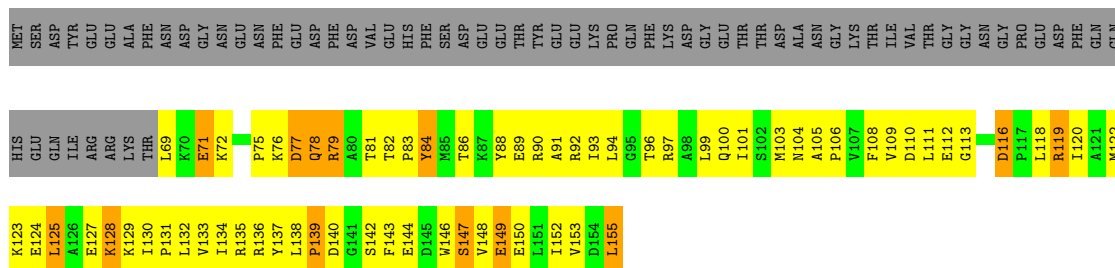
• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E: 20% 61% 17%



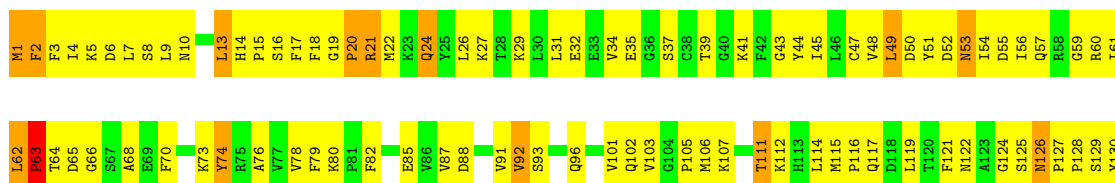
• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

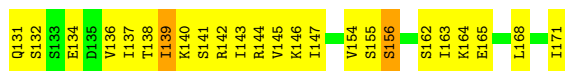
Chain F: 12% 35% 8% 44%



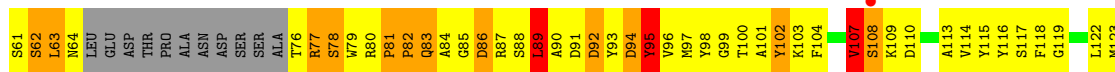
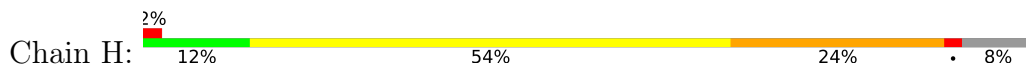
• Molecule 7: DNA-directed RNA polymerase II subunit RPB7

Chain G: 33% 58% 9%

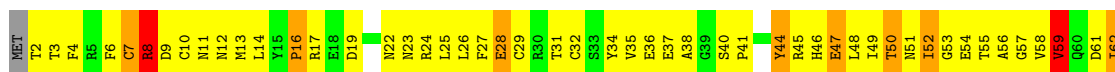




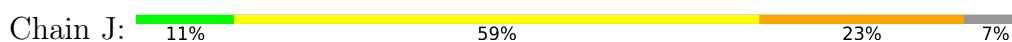
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3



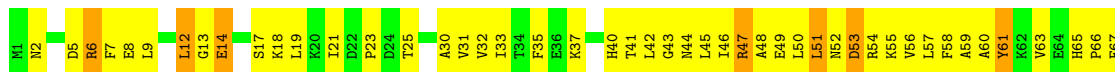
- Molecule 9: DNA-directed RNA polymerase II subunit RPB9



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

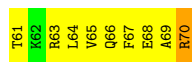


- Molecule 11: DNA-directed RNA polymerase II subunit RPB11



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





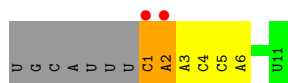
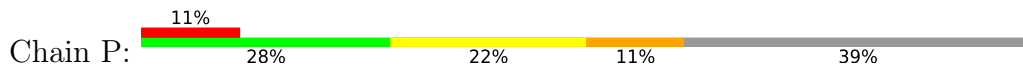
- Molecule 13: 5'-D(\*AP\*CP\*TP\*AP\*CP\*TP\*TP\*GP\*AP\*GP\*CP\*T)-3'



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



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



## 4 Data and refinement statistics i

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	221.62Å 393.72Å 282.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.65 49.21 – 3.65	Depositor EDS
% Data completeness (in resolution range)	99.9 (50.00-3.65) 100.0 (49.21-3.65)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.64 (at 3.67Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.212 , 0.250 0.214 , 0.253	Depositor DCC
$R_{free}$ test set	2686 reflections (1.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	86.1	Xtriage
Anisotropy	0.065	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 106.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.015 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.021 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	31918	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	73.0	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.*

<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: MG, ZN, BRU

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	0.53	1/11351 (0.0%)	0.79	6/15350 (0.0%)
2	B	0.52	0/9019	0.76	1/12160 (0.0%)
3	C	0.56	2/2133 (0.1%)	0.77	1/2891 (0.0%)
4	D	0.49	0/1444	0.76	2/1935 (0.1%)
5	E	0.50	0/1788	0.73	1/2406 (0.0%)
6	F	0.59	0/717	0.84	1/967 (0.1%)
7	G	0.55	0/1368	0.77	0/1844
8	H	0.54	0/1099	0.79	0/1488
9	I	0.49	0/989	0.75	0/1331
10	J	0.53	0/541	0.87	0/727
11	K	0.49	0/937	0.72	0/1265
12	L	0.61	0/365	0.82	0/485
13	N	0.96	0/152	1.05	0/232
14	T	0.84	0/391	0.84	0/600
15	P	0.70	0/237	0.86	0/368
All	All	0.54	3/32531 (0.0%)	0.78	12/44049 (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
14	T	0	3
All	All	0	5

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	88	CYS	CB-SG	-6.13	1.71	1.82

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	67	CYS	CB-SG	-5.53	1.72	1.81
3	C	92	CYS	CB-SG	-5.07	1.73	1.81

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	D	26	THR	N-CA-C	-6.91	92.33	111.00
1	A	56	PRO	N-CA-C	-6.77	94.50	112.10
1	A	1244	ARG	N-CA-C	6.05	127.33	111.00
3	C	39	ALA	N-CA-C	5.95	127.06	111.00
1	A	311	GLN	N-CA-C	5.68	126.34	111.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	1192	TYR	Sidechain
2	B	431	TYR	Sidechain
14	T	14	DA	Sidechain
14	T	16	DT	Sidechain
14	T	21	DC	Sidechain

## 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	11151	0	11226	1597	0
2	B	8847	0	8884	1363	0
3	C	2095	0	2051	348	0
4	D	1434	0	1460	218	0
5	E	1752	0	1776	259	0
6	F	705	0	731	113	0
7	G	1340	0	1357	180	0
8	H	1081	0	1051	207	0
9	I	971	0	929	152	0
10	J	532	0	542	118	0
11	K	919	0	929	116	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
12	L	363	0	388	97	0
13	N	137	0	82	10	0
14	T	369	0	202	27	0
15	P	213	0	110	19	0
16	A	2	0	0	0	0
16	B	1	0	0	0	0
16	C	1	0	0	0	0
16	I	2	0	0	0	0
16	J	1	0	0	0	0
16	L	1	0	0	0	0
17	A	1	0	0	0	0
All	All	31918	0	31718	4442	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 70.

The worst 5 of 4442 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:559:SER:HA	2:B:563:MET:HB3	1.20	1.16
12:L:55:ILE:HG12	12:L:56:LEU:H	0.97	1.12
12:L:55:ILE:CG1	12:L:56:LEU:H	1.59	1.12
3:C:43:THR:HG22	3:C:44:LEU:H	1.07	1.12
1:A:53:LEU:HD23	1:A:54:ASN:N	1.65	1.10

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	1407/1733 (81%)	977 (69%)	291 (21%)	139 (10%)	<b>0</b> <b>8</b>

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	1097/1224 (90%)	736 (67%)	235 (21%)	126 (12%)	0	5
3	C	264/347 (76%)	182 (69%)	50 (19%)	32 (12%)	0	5
4	D	174/221 (79%)	116 (67%)	41 (24%)	17 (10%)	0	8
5	E	212/215 (99%)	141 (66%)	51 (24%)	20 (9%)	0	9
6	F	85/155 (55%)	66 (78%)	15 (18%)	4 (5%)	2	22
7	G	169/171 (99%)	138 (82%)	23 (14%)	8 (5%)	2	22
8	H	131/146 (90%)	84 (64%)	21 (16%)	26 (20%)	0	1
9	I	117/122 (96%)	80 (68%)	26 (22%)	11 (9%)	0	9
10	J	63/70 (90%)	39 (62%)	12 (19%)	12 (19%)	0	1
11	K	112/120 (93%)	82 (73%)	25 (22%)	5 (4%)	2	23
12	L	44/70 (63%)	25 (57%)	7 (16%)	12 (27%)	0	0
All	All	3875/4594 (84%)	2666 (69%)	797 (21%)	412 (11%)	0	6

5 of 412 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	GLN
1	A	5	GLN
1	A	54	ASN
1	A	58	LEU
1	A	62	ASP

### 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	1240/1520 (82%)	1114 (90%)	126 (10%)	7	31
2	B	965/1061 (91%)	839 (87%)	126 (13%)	4	22
3	C	234/299 (78%)	203 (87%)	31 (13%)	4	21
4	D	160/200 (80%)	135 (84%)	25 (16%)	2	16
5	E	196/197 (100%)	176 (90%)	20 (10%)	7	31

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
6	F	77/137 (56%)	68 (88%)	9 (12%)	5	26
7	G	152/152 (100%)	138 (91%)	14 (9%)	9	36
8	H	118/128 (92%)	99 (84%)	19 (16%)	2	15
9	I	113/116 (97%)	97 (86%)	16 (14%)	3	19
10	J	60/65 (92%)	54 (90%)	6 (10%)	7	32
11	K	99/102 (97%)	91 (92%)	8 (8%)	11	41
12	L	40/57 (70%)	35 (88%)	5 (12%)	4	23
All	All	3454/4034 (86%)	3049 (88%)	405 (12%)	5	26

5 of 405 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	1058	LEU
4	D	17	LYS
11	K	114	LEU
2	B	1098	MET
3	C	73	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 116 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	465	ASN
8	H	131	ASN
2	B	984	HIS
8	H	128	ASN
5	E	101	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
15	P	10/18 (55%)	1 (10%)	1 (10%)

All (1) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
15	P	2	A

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
15	P	1	C

## 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).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
14	BRU	T	23	14,15	18,21,22	3.86	1 (5%)	26,30,33	1.01	2 (7%)

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
14	BRU	T	23	14,15	-	0/7/21/22	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
14	T	23	BRU	BR-C5	-16.27	1.50	1.88

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
14	T	23	BRU	C6-C5-C4	-2.86	117.77	120.67
14	T	23	BRU	C2'-C1'-N1	-2.04	109.06	113.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
14	T	23	BRU	1	0

## 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, 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	1417/1733 (81%)	-0.30	8 (0%) 89 83	18, 67, 100, 123	0
2	B	1113/1224 (90%)	-0.24	9 (0%) 86 77	21, 77, 109, 121	0
3	C	266/347 (76%)	-0.30	0 100 100	34, 66, 93, 107	0
4	D	178/221 (80%)	-0.14	0 100 100	46, 76, 106, 111	0
5	E	214/215 (99%)	-0.08	1 (0%) 91 85	43, 88, 108, 117	0
6	F	87/155 (56%)	-0.62	0 100 100	16, 44, 74, 84	0
7	G	171/171 (100%)	-0.28	0 100 100	47, 63, 94, 101	0
8	H	135/146 (92%)	0.26	3 (2%) 62 48	71, 94, 110, 116	0
9	I	119/122 (97%)	-0.00	3 (2%) 57 43	59, 92, 110, 119	0
10	J	65/70 (92%)	-0.41	0 100 100	45, 62, 87, 98	0
11	K	114/120 (95%)	-0.20	0 100 100	28, 67, 85, 96	0
12	L	46/70 (65%)	-0.04	0 100 100	51, 94, 110, 115	0
13	N	7/12 (58%)	2.06	4 (57%) 0 0	120, 125, 135, 137	0
14	T	17/26 (65%)	1.28	6 (35%) 0 0	99, 119, 135, 136	0
15	P	11/18 (61%)	1.12	2 (18%) 1 1	115, 119, 129, 139	0
All	All	3960/4650 (85%)	-0.22	36 (0%) 84 74	16, 73, 107, 139	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	471	LYS	5.8
15	P	1	C	3.9
2	B	470	LYS	3.7
9	I	116	ASN	3.3
1	A	155	GLU	3.3

## 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<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(Å <sup>2</sup> )	Q<0.9
14	BRU	T	23	20/21	0.73	0.29	105,113,118,121	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<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(Å <sup>2</sup> )	Q<0.9
17	MG	A	2458	1/1	0.83	0.15	151,151,151,151	0
16	ZN	L	1071	1/1	0.96	0.06	106,106,106,106	0
16	ZN	A	2456	1/1	0.96	0.06	87,87,87,87	0
16	ZN	I	1122	1/1	0.97	0.08	131,131,131,131	0
16	ZN	A	2457	1/1	0.98	0.17	42,42,42,42	0
16	ZN	B	2225	1/1	0.99	0.15	42,42,42,42	0
16	ZN	J	1066	1/1	0.99	0.20	47,47,47,47	0
16	ZN	C	1269	1/1	0.99	0.15	40,40,40,40	0
16	ZN	I	1121	1/1	0.99	0.09	79,79,79,79	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.