



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 09:16 AM GMT

PDB ID : 3HOV  
Title : Complete RNA polymerase II elongation complex II  
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 : 3.50 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](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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

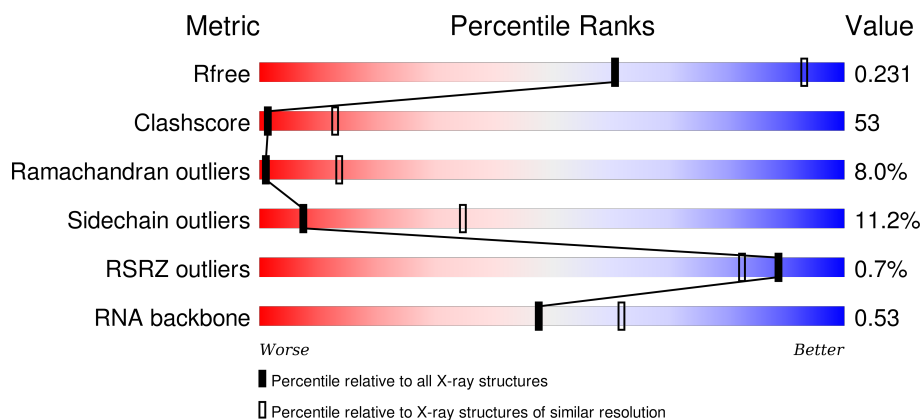
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	1051 (3.60-3.40)
Clashscore	102246	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)
RSRZ outliers	91569	1058 (3.60-3.40)
RNA backbone	2183	1050 (4.20-2.80)

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	A	1733	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 1%, yellow 44%, orange 9%, grey 18%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>27%</span> <span>44%</span> <span>9%</span> <span>18%</span> </div> </div>
2	B	1224	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red 1%, green 25%, yellow 52%, orange 12%, grey 10%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>25%</span> <span>52%</span> <span>12%</span> <span>10%</span> </div> </div>
3	C	318	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 26%, yellow 46%, orange 11%, grey 16%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>26%</span> <span>46%</span> <span>11%</span> <span>16%</span> </div> </div>
4	D	221	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green 25%, yellow 44%, orange 10%, red 1%, grey 19%);"></div> <div style="display: flex; justify-content: space-between; width: 90%; margin: 0 auto;"> <span>25%</span> <span>44%</span> <span>10%</span> <span>19%</span> </div> </div>

*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	T	26	
14	N	13	
15	P	17	

## 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	A	1416	Total	C	N	O	S	0	0	0
			11143	7021	1949	2111	62			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1104	Total	C	N	O	S	0	0	0
			8779	5560	1537	1627	55			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	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	D	178	Total	C	N	O	S	0	0	0
			1434	887	257	288	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	E	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	F	87	Total	C	N	O	S	0	0	0
			705	451	119	132	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	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	H	134	Total	C	N	O	S	0	0	0
			1076	677	182	213	4			

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

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

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
13	T	17	Total	Br	C	N	O	P	0	0	0
			347	1	166	61	103	16			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
14	N	5	Total	C	N	O	P	0	0	0
			97	48	18	27	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	P	10	Total	C	N	O	P	0	0	0
			215	96	42	67	10			

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

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

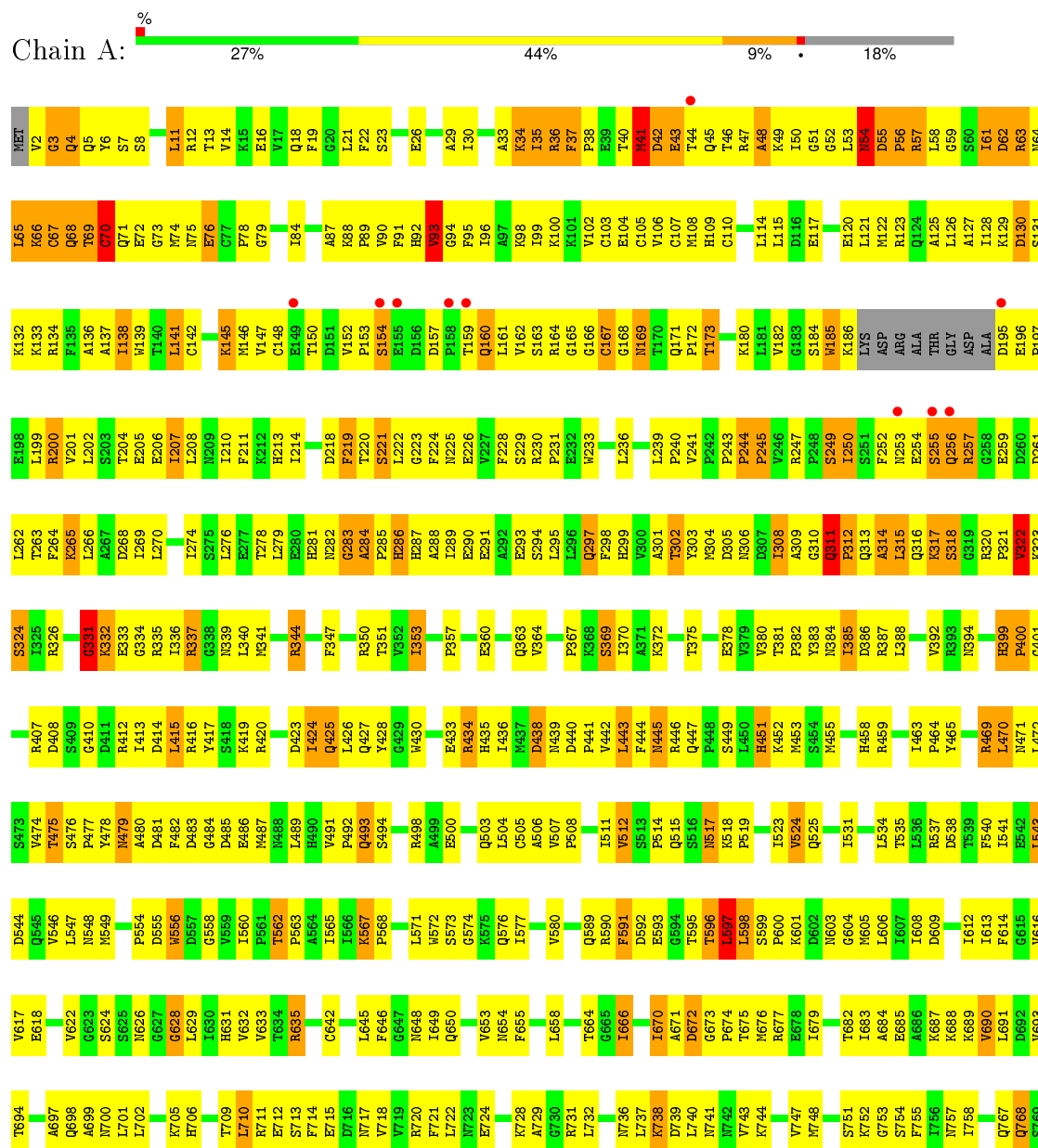
- Molecule 17 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	A	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 errors 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



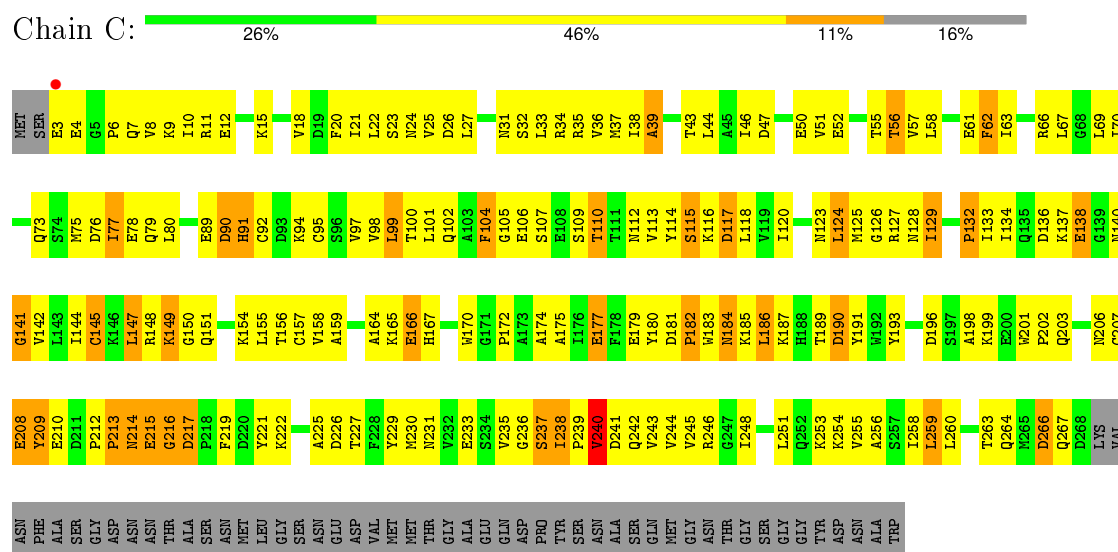




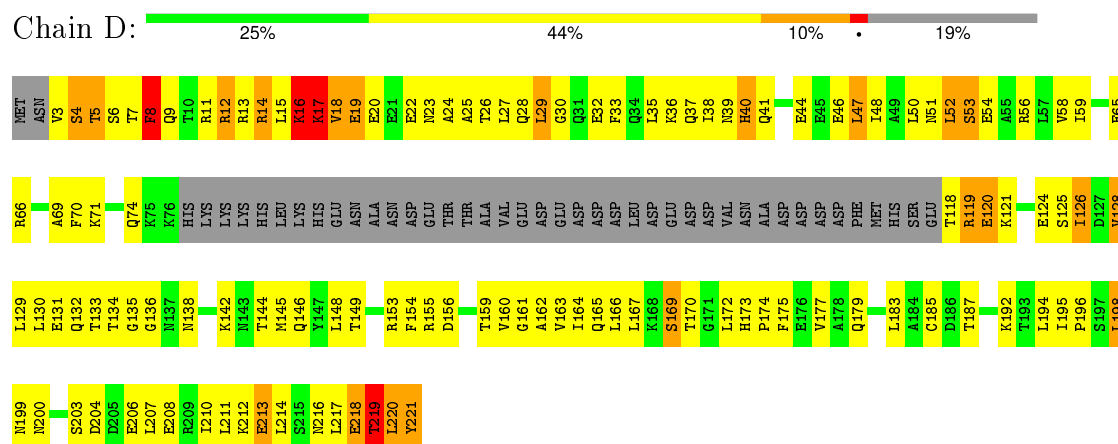
K985	P974	R983	S771	A704	D642	V580	B512	A447	F377	P316	G255	T185	T123	D61	ME1
V966	Q975	R984	A772	R705	D643	F581	Q513	I448	L378	C317	G256	E186	Y124	I62	SER
R969	Q976	S838	G773	Q706	E644	V582	L514	N449	G379	E319	K257	S187	S125	I63	ASP
T970	Q977	R839	G774	P707	S645	N583	B515	N450	V380	E318	L258	D188	S126	I64	LEU
T971	Q978	R840	K775	E708	L646	N586	N516	K451	K381	D320	G259	L189	E65	E65	ALA
	Q979	R841	K776	D709	G447	N586	T517	T452	I382		G260	Y190	L128	D66	ASN
		R842	A777	L710	H648		B518	I453	N383		G261	K191	L129	S67	SER
			M778	E711	K649	V589	B519	T454	K384		E262	L192	T68	T68	GLU
			G779	A713	E650	H590	G520	S454	I386		G263	K193	L69	L69	LYS
			V780	E714	K652	N592	L521	G457	L387		G264	E194	V130		TYR
				A715	H648		B526	K458			G265				
				ASN	K649	V589	B527	Y459			G266	M199	K134	LEU	GLY
				GLU	K650	V589	B528	A460	B391		G267	Y202	ARG	GLN	ASP
				GLU	K651	V589	B529	L461	B392		G268	F203	THR	LEU	ASP
				ASN	K652	V589	B530		K393		G269	F204	GLU	ALA	PRO
				ASP	K653	V589	B531	G464	D394		G270	I204	ALA	ALA	TYR
				LEU	K654	V589	B532	N465	Q395		G271	N206	ILE	HIS	GLY
				D722	K655	V589	B533	N466	D396		G272	N207	VAL	THR	PHE
				V723	K656	V589	B534	G467			G273	E209	PRO	THR	GLU
					K657	V589	B535	Q467	B399		G274	E210	GLY	GLU	D20
					K658	V589	B536	Q468	D399		G275	E211	GLY	SER	E21
					K659	V589	B537	Q469	H400		G276	V211	ARG	ASP	S22
					K660	V589	B538	K470	F401		G277		GLU	ASN	A23
					K661	V589	B539	L471	Q402		G278	E212	LEU	ILE	A23
					K662	V589	B540	L472	Q403		G279	E213	LEU	ILE	P24
					K663	V589	B541	L473	Q404		G280	E214	LEU	ILE	I25
					K664	V589	B542	L474	Q405		G281	E215	SER	SER	I26
					K665	V589	B543	L475	Q406		G282	E216	ARG	ARG	A27
					K666	V589	B544	L476	Q407		G283	E217	GLU	LYS	E28
					K667	V589	B545	L477	Q408		G284	E218	GLU	TYR	D29
					K668	V589	B546	L478	Q409		G285	E219	GLU	ILE	D30
					K669	V589	B547	L479	Q410		G286	E220	GLU	S91	S30
					K670	V589	B548	L480	Q411		G287	E221	GLU	F92	K31
					K671	V589	B549	L481	Q412		G288	E222	GLU	F93	K32
					K672	V589	B550	L482	Q413		G289	E223	GLU	F94	V33
					K673	V589	B551	L483	Q414		G290	E224	GLU	F95	S35
					K674	V589	B552	L484	Q415		G291	E225	GLU	F96	A36
					K675	V589	B553	L485	Q416		G292	E226	GLU	F97	F37
					K676	V589	B554	L486	Q417		G293	E227	GLU	F98	F38
					K677	V589	B555	L487	Q418		G294	E228	GLU	F99	F39
					K678	V589	B556	L488	Q419		G295	E229	GLU	P100	E40
					K679	V589	B557	L489	Q420		G296	E230	GLU	M101	K41
					K680	V589	B558	L490	Q421		G297	E231	GLU	V102	G42
					K681	V589	B559	L491	Q422		G298	E232	GLU	N103	L43
					K682	V589	B560	L492	Q423		G299	E233	GLU	E104	V44
					K683	V589	B561	L493	Q424		G300	E234	GLU	S105	S45
					K684	V589	B562	L494	Q425		G301	E235	GLU	D106	Q46
					K685	V589	B563	L495	Q426		G302	E236	GLU	G107	Q47
					K686	V589	B564	L496	Q427		G303	E237	GLU	V108	L48
					K687	V589	B565	L497	Q428		G304	E238	GLU	T109	D49
					K688	V589	B566	L498	Q429		G305	E239	GLU	H110	S50
					K689	V589	B567	L499	Q430		G306	E240	GLU	M111	F51
					K690	V589	B568	L500	Q431		G307	E241	GLU	L112	N52
					K691	V589	B569	L501	Q432		G308	E242	GLU	Y113	Q53
					K692	V589	B570	L502	Q433		G309	E243	GLU	P114	F54
					K693	V589	B571	L503	Q434		G310	E244	GLU	Q115	V55
					K694	V589	B572	L504	Q435		G311	E245	GLU	R116	D66
					K695	V589	B573	L505	Q436		G312	E246	GLU	Y117	Y57
					K696	V589	B574	L506	Q437		G313	E247	GLU	L118	T58
					K697	V589	B575	L507	Q438		G314	E248	GLU	R119	L59
					K698	V589	B576	L508	Q439		G315	E249	GLU	L120	L60
					K699	V589	B577	L509	Q440		G316	E250	GLU	S182	Q60
					K700	V589	B578	L510	Q441		G317	E251	GLU	L254	
					K701	V589	B579	L511	Q442		G318	E252	GLU	L255	
					K702	V589	B580	L512	Q443		G319	E253	GLU	L256	
					K703	V589	B581	L513	Q444		G320	E254	GLU	L257	
					K704	V589	B582	L514	Q445		G321	E255	GLU	L258	
					K705	V589	B583	L515	Q446		G322	E256	GLU	L259	
					K706	V589	B584	L516	Q447		G323	E257	GLU	L260	
					K707	V589	B585	L517	Q448		G324	E258	GLU	L261	
					K708	V589	B586	L518	Q449		G325	E259	GLU	L262	
					K709	V589	B587	L519	Q450		G326	E260	GLU	L263	
					K710	V589	B588	L520	Q451		G327	E261	GLU	L264	
					K711	V589	B589	L521	Q452		G328	E262	GLU	L265	
					K712	V589	B590	L522	Q453		G329	E263	GLU	L266	
					K713	V589	B591	L523	Q454		G330	E264	GLU	L267	
					K714	V589	B592	L524	Q455		G331	E265	GLU	L268	
					K715	V589	B593	L525	Q456		G332	E266	GLU	L269	
					K716	V589	B594	L526	Q457		G333	E267	GLU	L270	
					K717	V589	B595	L527	Q458		G334	E268	GLU	L271	
					K718	V589	B596	L528	Q459		G335	E269	GLU	L272	
					K719	V589	B597	L529	Q460		G336	E270	GLU	L273	
					K720	V589	B598	L530	Q461		G337	E271	GLU	L274	
					K721	V589	B599	L531	Q462		G338	E272	GLU	L275	
					K722	V589	B600	L532	Q463		G339	E273	GLU	L276	
					K723	V589	B601	L533	Q464		G340	E274	GLU	L277	
					K724	V589	B602	L534	Q465		G341	E275	GLU	L278	
					K725	V589	B603	L535	Q466		G342	E276	GLU	L279	
					K726	V589	B604	L536	Q467		G343	E277	GLU	L280	
					K727	V589	B605	L537	Q468		G344	E278	GLU	L281	
					K728	V589	B606	L538	Q469		G345	E279	GLU	L282	
					K729	V589	B607	L539	Q470		G346	E280	GLU	L283	
					K730	V589	B608	L540	Q471		G347	E281	GLU	L284	
					K731	V589	B609	L541	Q472		G348	E282	GLU	L285	
					K732	V589	B610	L542	Q473		G349	E283	GLU	L286	
					K733	V589	B611	L543	Q474		G350	E284	GLU	L287	
					K734	V589	B612	L544	Q475		G351	E285	GLU	L288	
					K735	V589	B613	L545	Q476		G352	E286	GLU	L289	
					K736	V589	B614	L546	Q477		G353	E287	GLU	L290	
					K737	V589	B615	L547	Q478		G354	E288	GLU	L291	
					K738	V589	B616	L548	Q479		G355	E289	GLU	L292	
					K739	V589	B617	L549	Q480		G356	E290	GLU	L293	
					K740	V589	B618	L550	Q481		G357	E291	GLU	L294	
					K741	V589	B619	L551	Q482		G358	E292	GLU	L295	
					K742	V589	B620	L552	Q483		G359	E293	GLU	L296	
					K743	V589	B621	L553	Q484		G360	E294	GLU	L297	
					K744	V589	B622	L554	Q485		G361	E295	GLU	L298	
					K745	V589	B623	L555	Q486		G362	E296	GLU	L299	
					K746	V589	B624	L556	Q487		G363	E297	GLU	L300	
					K747	V589	B625	L557	Q488		G364	E298	GLU	L301	
					K748	V589	B626	L558	Q489		G365	E299	GLU	L302	
					K749	V589	B627	L559	Q490		G366	E300	GLU	L303	
					K750	V589	B628	L560	Q491		G367	E301	GLU	L304	
					K751	V589	B629	L561	Q492		G368	E302	GLU	L305	
					K752	V589	B630	L562	Q493		G369	E303	GLU	L306	



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



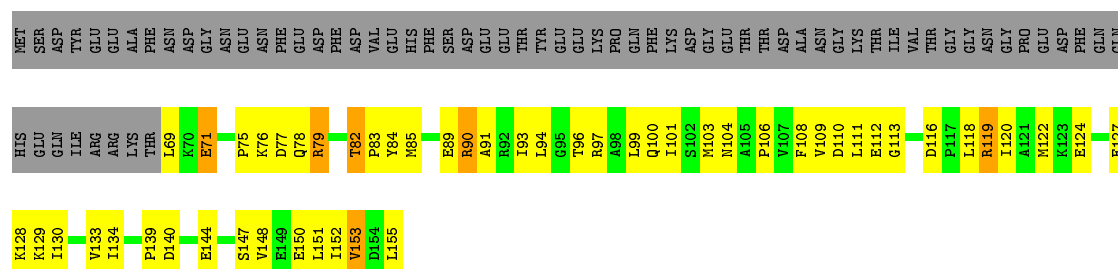
• Molecule 4: DNA-directed RNA polymerase II subunit RPB4





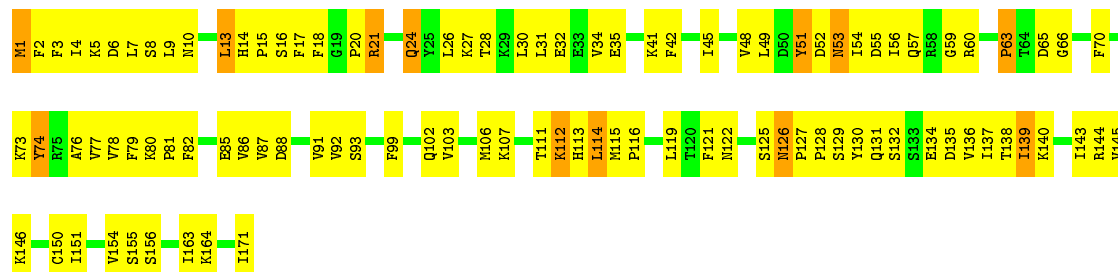
- Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F: 23% 30% 44%



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

Chain G: 40% 53% 7%



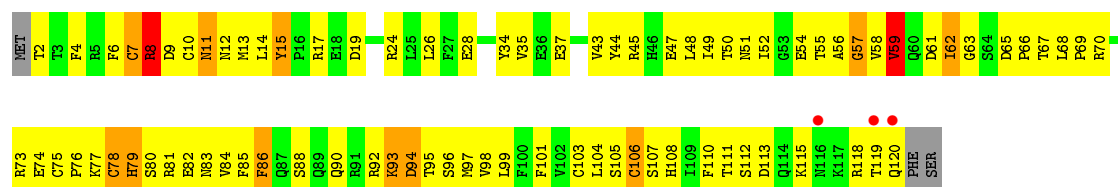
- Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H: 18% 57% 15% 8%

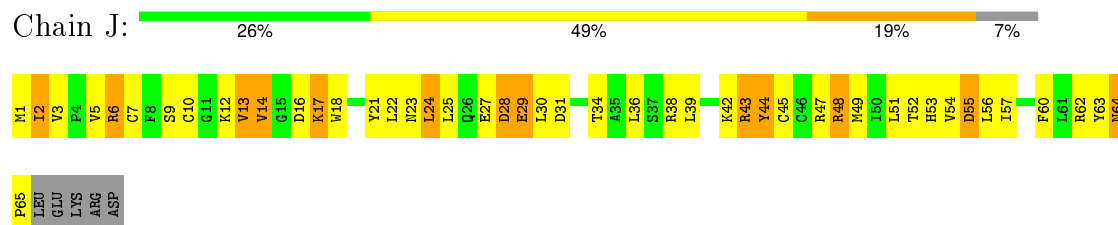


- Molecule 9: DNA-directed RNA polymerase II subunit RPB9

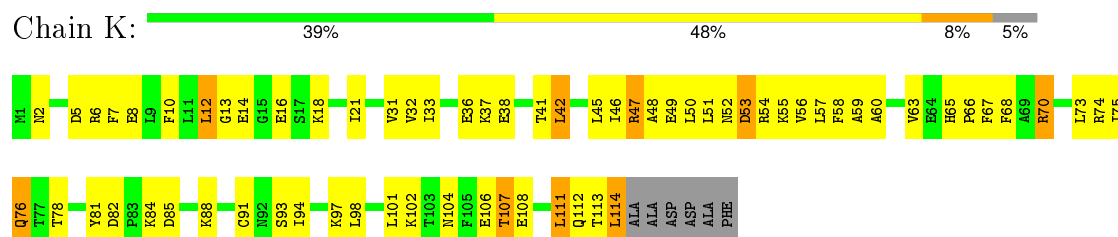
Chain I: 30% 57% 9%



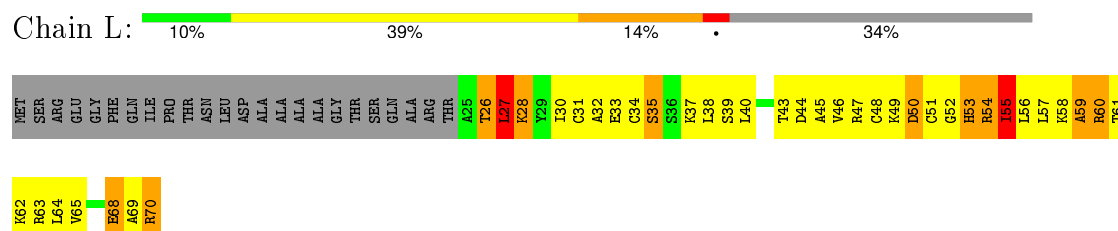
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5



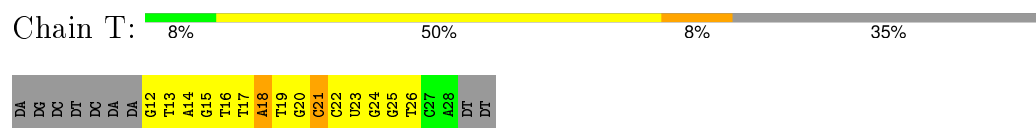
- Molecule 11: DNA-directed RNA polymerase II subunit RPB11



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



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



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



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	222.25Å 393.38Å 283.09Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.50 49.17 – 3.50	Depositor EDS
% Data completeness (in resolution range)	99.9 (50.00-3.50) 99.9 (49.17-3.50)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.76 (at 3.48Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.210 , 0.226 0.212 , 0.231	Depositor DCC
$R_{free}$ test set	4056 reflections (2.68%)	DCC
Wilson B-factor (Å <sup>2</sup> )	85.9	Xtriage
Anisotropy	0.056	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 99.2	EDS
Estimated twinning fraction	0.017 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
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtriage
Outliers	1 of 155149 reflections (0.001%)	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	31777	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	82.0	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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.49	0/11342	0.77	12/15337 (0.1%)
2	B	0.48	0/8948	0.74	1/12062 (0.0%)
3	C	0.47	0/2133	0.75	1/2891 (0.0%)
4	D	0.45	0/1444	0.72	1/1935 (0.1%)
5	E	0.45	0/1788	0.68	1/2406 (0.0%)
6	F	0.56	0/717	0.80	1/967 (0.1%)
7	G	0.48	0/1368	0.76	1/1844 (0.1%)
8	H	0.43	0/1094	0.72	0/1481
9	I	0.44	0/989	0.71	0/1331
10	J	0.50	0/541	0.83	0/727
11	K	0.48	0/937	0.67	0/1265
12	L	0.57	0/365	0.84	0/485
13	T	0.54	0/365	0.85	0/560
14	N	0.57	0/108	0.78	0/164
15	P	0.45	0/240	0.82	1/372 (0.3%)
All	All	0.48	0/32379	0.75	19/43827 (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	1
8	H	0	1
13	T	0	3
All	All	0	5

There are no bond length outliers.

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	56	PRO	N-CA-C	-6.45	95.34	112.10
3	C	39	ALA	N-CA-C	6.38	128.24	111.00
1	A	55	ASP	N-CA-CB	6.13	121.63	110.60
1	A	3	GLY	N-CA-C	-6.04	97.99	113.10
15	P	1	C	N1-C1'-C2'	5.90	121.67	114.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	1192	TYR	Sidechain
8	H	102	TYR	Sidechain
13	T	18	DA	Sidechain
13	T	19	DT	Sidechain
13	T	21	DC	Sidechain

## 5.2 Too-close contacts

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	11143	0	11217	1194	0
2	B	8779	0	8808	1069	0
3	C	2095	0	2051	259	0
4	D	1434	0	1460	156	0
5	E	1752	0	1776	193	0
6	F	705	0	731	75	0
7	G	1340	0	1357	133	0
8	H	1076	0	1046	159	0
9	I	971	0	929	118	0
10	J	532	0	542	77	0
11	K	919	0	929	82	0
12	L	363	0	388	85	0
13	T	347	0	192	20	0
14	N	97	0	58	6	0
15	P	215	0	111	7	0
16	A	2	0	0	0	0
16	B	1	0	0	0	0
16	C	1	0	0	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
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	31777	0	31595	3348	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 53.

The worst 5 of 3348 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	1.18
2:B:583:ASN:HD21	2:B:628:THR:HG22	1.08	1.17
2:B:622:LYS:HE2	9:I:59:VAL:HG22	1.28	1.14
2:B:261:ARG:HH11	2:B:261:ARG:HB3	1.12	1.13
1:A:53:LEU:HD23	1:A:54:ASN:N	1.65	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	A	1406/1733 (81%)	1062 (76%)	240 (17%)	104 (7%)	1	16
2	B	1082/1224 (88%)	795 (74%)	188 (17%)	99 (9%)	1	11
3	C	264/318 (83%)	200 (76%)	45 (17%)	19 (7%)	1	17
4	D	174/221 (79%)	131 (75%)	29 (17%)	14 (8%)	1	13
5	E	212/215 (99%)	159 (75%)	40 (19%)	13 (6%)	2	21
6	F	85/155 (55%)	72 (85%)	12 (14%)	1 (1%)	16	61

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
7	G	169/171 (99%)	143 (85%)	20 (12%)	6 (4%)	4	37
8	H	130/146 (89%)	86 (66%)	23 (18%)	21 (16%)	0	3
9	I	117/122 (96%)	76 (65%)	31 (26%)	10 (8%)	1	13
10	J	63/70 (90%)	44 (70%)	10 (16%)	9 (14%)	0	4
11	K	112/120 (93%)	90 (80%)	19 (17%)	3 (3%)	6	44
12	L	44/70 (63%)	18 (41%)	16 (36%)	10 (23%)	0	1
All	All	3858/4565 (84%)	2876 (74%)	673 (17%)	309 (8%)	1	13

5 of 309 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	GLN
1	A	43	GLU
1	A	54	ASN
1	A	57	ARG
1	A	58	LEU

### 5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	1239/1520 (82%)	1099 (89%)	140 (11%)	7	34
2	B	958/1061 (90%)	845 (88%)	113 (12%)	6	31
3	C	234/274 (85%)	205 (88%)	29 (12%)	6	29
4	D	160/200 (80%)	134 (84%)	26 (16%)	3	17
5	E	196/197 (100%)	180 (92%)	16 (8%)	14	51
6	F	77/137 (56%)	70 (91%)	7 (9%)	12	46
7	G	152/152 (100%)	138 (91%)	14 (9%)	11	45
8	H	118/128 (92%)	105 (89%)	13 (11%)	8	36
9	I	113/116 (97%)	106 (94%)	7 (6%)	23	63
10	J	60/65 (92%)	53 (88%)	7 (12%)	7	32

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
11	K	99/102 (97%)	91 (92%)	8 (8%)	15	52
12	L	40/57 (70%)	35 (88%)	5 (12%)	6	29
All	All	3446/4009 (86%)	3061 (89%)	385 (11%)	7	35

5 of 385 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	452	THR
2	B	889	THR
8	H	138	GLU
2	B	475	SER
2	B	645	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 100 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	484	ASN
2	B	835	GLN
8	H	128	ASN
2	B	513	GLN
2	B	573	GLN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
15	P	9/17 (52%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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

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	$\# Z  > 2$
13	BRU	T	23	13,15	13,21,22	4.59	4 (30%)	16,30,33	4.04	3 (18%)

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
13	BRU	T	23	13,15	-	0/3/21/22	0/2/2/2

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
13	T	23	BRU	BR-C5	-15.08	1.50	1.90
13	T	23	BRU	C6-N1	2.57	1.38	1.35
13	T	23	BRU	C4-N3	3.31	1.39	1.33
13	T	23	BRU	C4-C5	5.13	1.45	1.38

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
13	T	23	BRU	C5-C4-N3	-8.12	115.33	124.00
13	T	23	BRU	C5-C6-N1	2.05	123.81	119.79
13	T	23	BRU	C4-N3-C2	13.61	127.01	115.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
13	T	23	BRU	3	0

## 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	A	1416/1733 (81%)	-0.24	11 (0%) 87 80	28, 72, 115, 145	0
2	B	1104/1224 (90%)	-0.19	8 (0%) 89 82	28, 85, 126, 140	0
3	C	266/318 (83%)	-0.21	1 (0%) 93 90	41, 71, 107, 125	0
4	D	178/221 (80%)	-0.08	0 100 100	52, 84, 120, 134	0
5	E	214/215 (99%)	-0.02	0 100 100	52, 103, 130, 138	0
6	F	87/155 (56%)	-0.53	0 100 100	32, 51, 81, 91	0
7	G	171/171 (100%)	-0.21	0 100 100	48, 71, 104, 117	0
8	H	134/146 (91%)	0.26	2 (1%) 76 67	79, 109, 127, 139	0
9	I	119/122 (97%)	0.06	3 (2%) 61 50	66, 106, 126, 144	0
10	J	65/70 (92%)	-0.36	0 100 100	49, 68, 95, 107	0
11	K	114/120 (95%)	-0.21	0 100 100	36, 76, 95, 112	0
12	L	46/70 (65%)	0.03	0 100 100	48, 111, 129, 130	0
13	T	16/26 (61%)	0.50	0 100 100	99, 137, 166, 166	0
14	N	5/13 (38%)	1.43	1 (20%) 1 2	143, 147, 156, 171	0
15	P	10/17 (58%)	0.66	1 (10%) 9 9	122, 129, 147, 153	0
All	All	3945/4621 (85%)	-0.18	27 (0%) 89 82	28, 80, 124, 171	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1455	PRO	4.0
2	B	918	ILE	3.2
9	I	120	GLN	3.0
14	N	1	DA	2.9
15	P	1	C	2.9

## 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. 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(Å <sup>2</sup> )	Q<0.9
13	BRU	T	23	20/21	0.74	0.26	-	127,131,137,139	0

## 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(Å <sup>2</sup> )	Q<0.9
16	ZN	J	1066	1/1	0.99	0.21	-0.37	61,61,61,61	0
16	ZN	B	2225	1/1	0.99	0.17	-1.07	54,54,54,54	0
16	ZN	C	1269	1/1	0.99	0.10	-1.46	49,49,49,49	0
16	ZN	A	2457	1/1	0.99	0.14	-1.82	51,51,51,51	0
16	ZN	I	1121	1/1	0.99	0.10	-1.92	90,90,90,90	0
16	ZN	L	1071	1/1	0.91	0.08	-2.25	111,111,111,111	0
16	ZN	I	1122	1/1	0.95	0.05	-3.00	122,122,122,122	0
16	ZN	A	2456	1/1	0.97	0.06	-4.45	86,86,86,86	0
17	MG	A	2458	1/1	0.73	0.32	-	123,123,123,123	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.