



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 08:01 PM GMT

PDB ID : 1I3Q  
Title : RNA POLYMERASE II CRYSTAL FORM I AT 3.1 Å RESOLUTION  
Authors : Cramer, P.; Bushnell, D.A.; Kornberg, R.D.  
Deposited on : 2001-02-15  
Resolution : 3.10 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) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

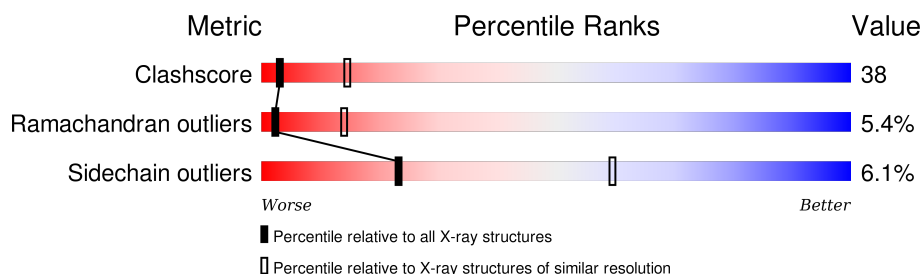
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.10 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))
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)




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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	1733	
2	B	1224	
3	C	318	
4	E	215	
5	F	155	
6	H	146	
7	I	122	

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
8	J	70	
9	K	120	
10	L	70	

## 2 Entry composition

There are 12 unique types of molecules in this entry. The entry contains 28161 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 LARGEST SUB-UNIT.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1414	Total	C	N	O	S	0	0	0
			11114	7000	1947	2106	61			

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE II 140KD POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1083	Total	C	N	O	S	0	0	0
			8624	5470	1501	1600	53			

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE II 45KD POLYPEPTIDE.

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 27KD POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	215	Total	C	N	O	S	0	0	0
			1760	1116	310	322	12			

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASE II 23KD POLYPEPTIDE.

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

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASE II 14.5KD POLYPEPTIDE.

TIDE.

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

- Molecule 7 is a protein called DNA-DIRECTED RNA POLYMERASE II 14.2KD POLYPEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	I	122	Total	C	N	O	S	0	0	0
			997	613	182	191	11			

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASE II 8.3KD POLYPEPTIDE.

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

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II 13.6KD POLYPEPTIDE.

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

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASE II 7.7KD POLYPEPTIDE.

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

- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	J	1	Total	Zn	0	0
			1	1		
11	B	1	Total	Zn	0	0
			1	1		
11	I	2	Total	Zn	0	0
			2	2		
11	C	1	Total	Zn	0	0
			1	1		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	A	2	Total 2	Zn 2	0	0
11	L	1	Total 1	Zn 1	0	0

- Molecule 12 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

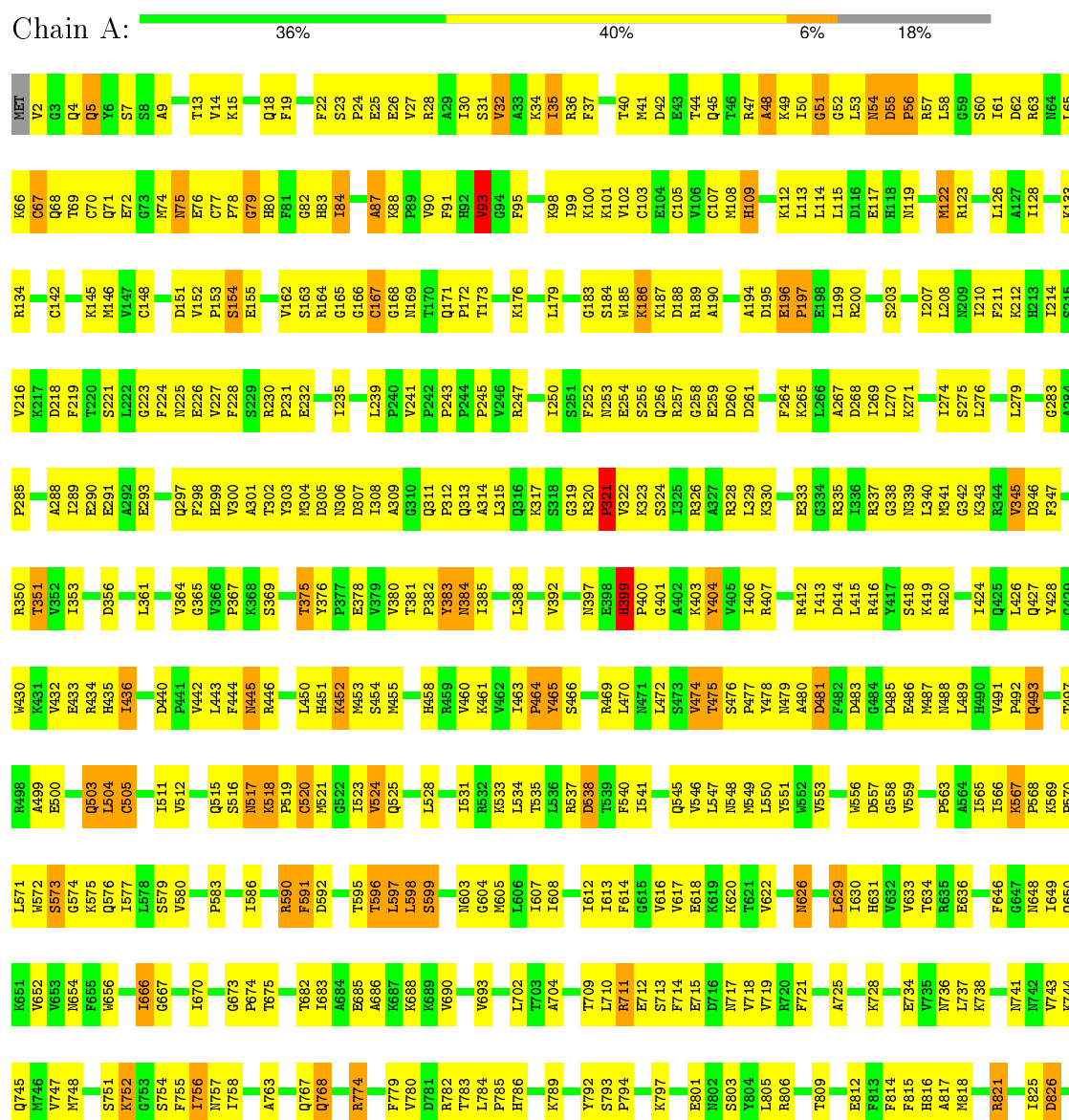
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
12	A	1	Total 1	Mg 1	0	0

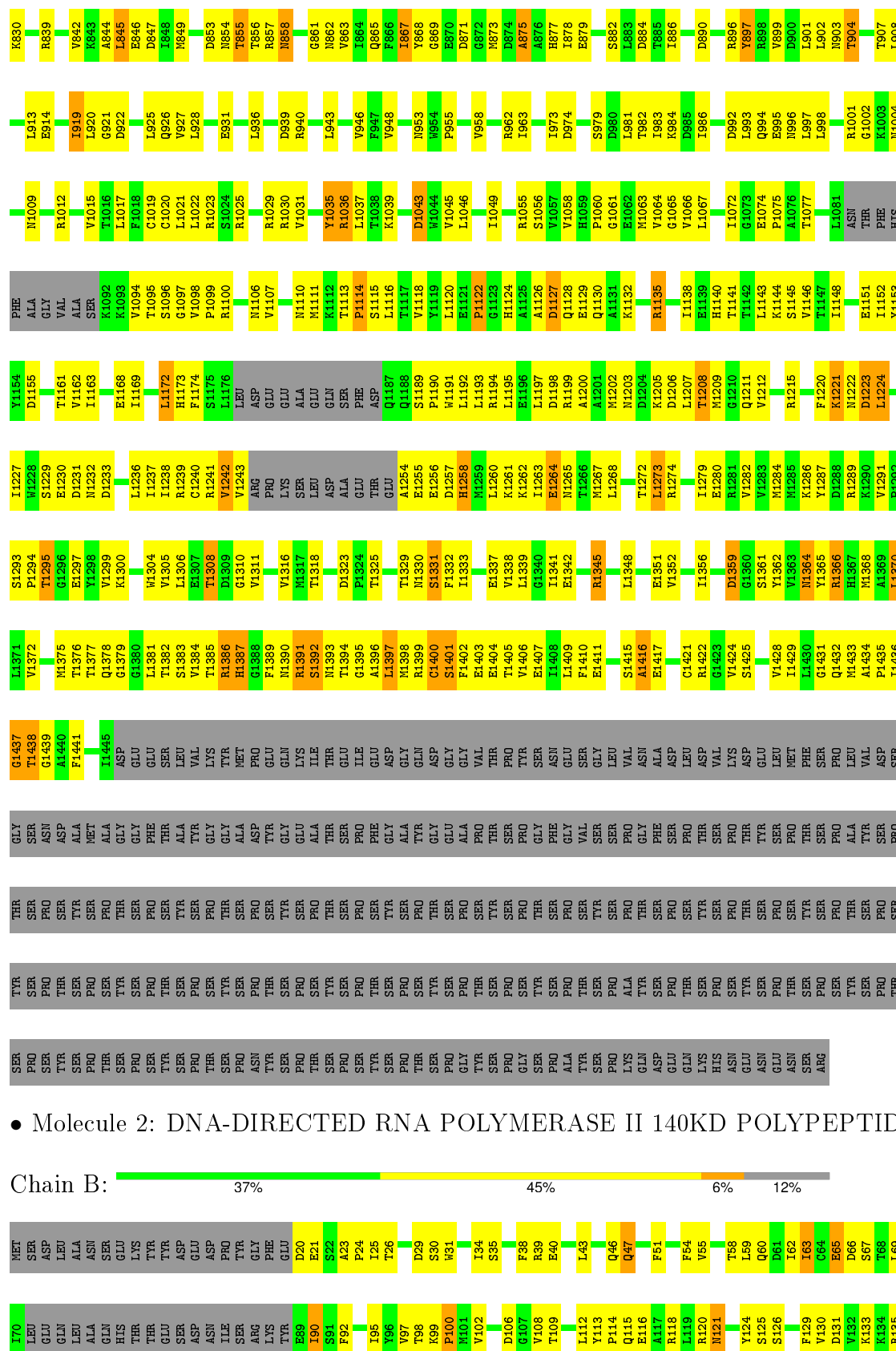
### 3 Residue-property plots

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.

Note EDS was not executed.

#### • Molecule 1: DNA-DIRECTED RNA POLYMERASE II LARGEST SUBUNIT







R1222	F1146	T1077	K934	Q862	A793	GLU	R654	B579	LVS	E437	Y351	L282	Y202	T136
D1223	L1147	K1078	R935	E863	N794	GLU	K655	V580	LEU	GLU	A352	V283	F203	Y137
F1224	K1148	K1079	D936	K864	L796	ASN	G656	F581	A509	ALA	K353	L284	I204	E138
	M1152	Q1084	S938	T871	Y797	LEU	I658	V585	K501	HIS		L285	I205	ALA
	A1153	E1004	I941	E872	Y798		I659	V586	P511	ASP	F360	R287	N206	I1E
	A1154	G1005	R942	T873	P799		K660	V587	R512	PHE	F361	A288	G207	ASP
	S1155	I1006	S943	F874	Q800		L661	G588	Q513	ASN	L361		S208	VAL
	D1156	F1087			R801	A726	P662	V589	H515	MET	P362	E209	E209	PRO
	R1094	R1008	T944	R879	P802		T664	B590	N516	L446	G363	L291	L212	GLY
	F1158	E345	R945	T880	T805	R729	E665	L596	T517	A447	T365	P293	L211	ARG
	R1159	R946	G947	N881	T806	S731	V666	V597	H518	L448	G366	L292	L212	GLU
	V1160	T882		T882	R807	R732	Q667	E598	H519		L367	G295	E216	LEU
	H1161	I948		L883	A808	H734	D668	T599	G520	N449		E296		LVS
	I1162	V952		R884	M809	A735	I1E	V521		A450	F370	L297		TTR
		P1014		H887	R808		GLU	L600	V522			L298	G220	GLU
		H1015	L953	G887	E810	T736	GLY	B601	L521	L453	R373	L299	N221	LEU
		A1016	V954	G888	Y811	T737	GLY	T502	Q531	T454	K374	E299		ILE
		I1017	T955	T889	L812	T738	GLY	L603	A532	S455	A375	H300	V225	ALA
		P1018	T956	Y890	K813	T739	PHE	R604	G553		F376		F226	GLU
		S1019	R957	D891	F814	H740	GLU	R605	G554	L461	F377	V305	K227	GLU
		R1020	Q958	K892	R815	H743	ASP		L535	A462	L378	R306	K228	SER
		M1021	D959	L893	E816	T744	VAL	L609	V536	T463		D307	A229	GLU
		T1022	G960	K894	L817	R745	GLU		K537	T464	M381	V308	A230	ASP
		H1025	R961	D896	R818	R746	E678	B612	M538	N466	I382	L311	P231	ASP
		E1028	G962	G897	A819	S746	V679	V613		G467				SER
		G1029	R963	L898	G820	M747	T680		L541	GLU	L387	L314	I234	GLU
		L1030	V966	R900	Q821	T748	W681	1616	M542	GLN		K315	S235	SER
		V1034	R969	A901	A823	V751	S682	R617		LVS	D391	P316	H236	GLY
		A1035	T970	G902	I824	T755	L684	D618	I545	ALA	R392	C317	V237	V165
		A1036	T971	Y905	W825	T756	L685	1619	S546	ALA	K393	D320	I240	F166
		L1037	R972	G906	A826	T757	L686	R620	V547	MET	D394	G321	R241	I167
		E1041	P973	G907	I827	T758	E687	B521	G548	SER	Q395	G322	S242	G168
		G1042	Q975	I911	W830	P759	L689	L624	T549	ARG	D396	V323	A243	L170
		D1043	T976	I912			V690		P551	A477	R398	L324	I244	P171
		A1044	G977	G913	Q835	Q763	E691	T627	M552	S483	G402	Q325	G247	I172
		S1045	D978	K914	I840	S764	F692	D629	P553			D326	S248	H173
		P1046	K979	T915	N841	P766	D694	A630	I554	L483	R405	T329	R249	L174
		I1050	F980	T916	N842	T767	A695	R632	F556	N484	L408	A330	I251	N178
		T1051	A981	P917	Q843	T768	E696	V633	F557	R485	A409	D332	S252	G179
		V1052	S982	I918	S844	T769	E697	R634	L558	T487	F417	F333	L254	Y180
		E1053	R983	S919	I846	Q770		R635	S559	Y488	R418	I334		L181
		G1054	Q985	ASP	D847	M778	L702	P636	G562	S489		G335	Y259	T185
		S1056	K987	GLU	L850	G779	T701	L637	E564	S490	F421	ARG	R261	T189
		R1060	G988	GLU	F851	T781	M705	F638	P565	B494	L424	THR	R267	Y190
		D1136	T989	LEU	L854	L782	Q706	V640	E564	L495	T425	ALA		
		M1133	T990	GLY	F855	M784	P707	B641	Y569	R496				
		F1139	G991	GLN	R857	T785	E708	D642	V570	R497	I428	LEU	L273	K193
		A1140	I992	ARG	R857	T787	L710	B644		N499	F429	GLY	P274	E194
		R1067	R985	THR	S858	R788	E711		Q573		R430	LVS	I276	C195
		A1144	R996	ALA	R859	H789	P712	B649	P575	L502	Y431	K345	K277	F197
		S1145	T997	HIS	M860	T790	ALA	B650	P576	GLY	M432	E346	Q278	D198
			D998			M792	GLU	L651	D576	ARG	T435	R347	I279	M199
							ASN	V653	T578			R348	I280	G200
													P281	G201

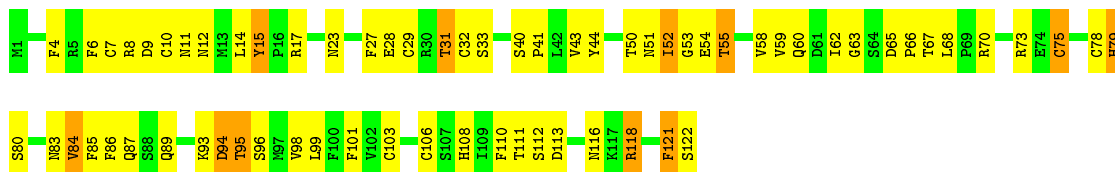
- Molecule 3: DNA-DIRECTED RNA POLYMERASE II 45KD POLYPEPTIDE





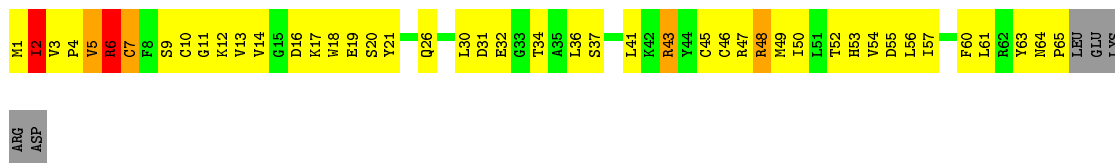
• Molecule 7: DNA-DIRECTED RNA POLYMERASE II 14.2KD POLYPEPTIDE

Chain I: 45% 46% 9%



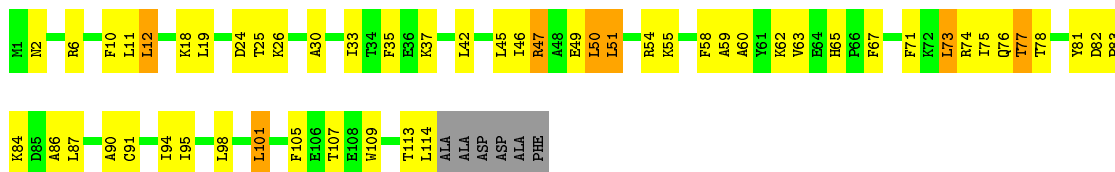
• Molecule 8: DNA-DIRECTED RNA POLYMERASE II 8.3KD POLYPEPTIDE

Chain J: 29% 56% 6% 7%



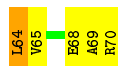
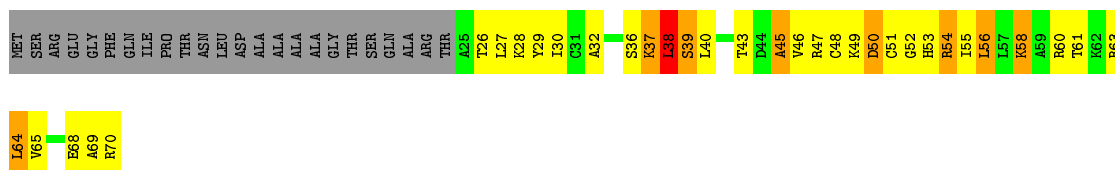
• Molecule 9: DNA-DIRECTED RNA POLYMERASE II 13.6KD POLYPEPTIDE

Chain K: 50% 39% 6% 5%



• Molecule 10: DNA-DIRECTED RNA POLYMERASE II 7.7KD POLYPEPTIDE

Chain L: 19% 34% 11% 34%



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	130.70 Å   224.80 Å   369.40 Å 90.00°   90.00°   90.00°	Depositor
Resolution (Å)	40.00 – 3.10	Depositor
% Data completeness (in resolution range)	(Not available) (40.00-3.10)	Depositor
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
Refinement program	CNS	Depositor
R, $R_{free}$	0.229 , 0.283	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	28161	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	71.0	wwPDB-VP

## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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	A	0.40	1/11312 (0.0%)	0.70	2/15298 (0.0%)
2	B	0.40	0/8793	0.68	3/11857 (0.0%)
3	C	0.42	0/2133	0.72	1/2891 (0.0%)
4	E	0.37	0/1796	0.63	0/2416
5	F	0.42	0/691	0.66	0/933
6	H	0.59	2/1086 (0.2%)	0.91	6/1470 (0.4%)
7	I	0.41	0/1016	0.70	0/1365
8	J	0.43	0/541	0.85	1/727 (0.1%)
9	K	0.42	0/937	0.61	0/1265
10	L	0.41	0/366	0.66	0/485
All	All	0.41	3/28671 (0.0%)	0.70	13/38707 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	H	109	LYS	CD-CE	7.54	1.70	1.51
6	H	109	LYS	CE-NZ	5.69	1.63	1.49
1	A	520	CYS	CB-SG	-5.67	1.72	1.81

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	H	109	LYS	N-CA-C	7.39	130.95	111.00
1	A	452	LYS	N-CA-C	-6.74	92.80	111.00
6	H	109	LYS	CA-CB-CG	6.68	128.09	113.40
6	H	80	ARG	NE-CZ-NH1	-6.09	117.25	120.30
6	H	80	ARG	NE-CZ-NH2	5.97	123.28	120.30

There are no chirality outliers.

There are no planarity outliers.

## 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	11114	0	11193	945	0
2	B	8624	0	8642	736	1
3	C	2095	0	2051	177	0
4	E	1760	0	1788	103	0
5	F	679	0	701	56	0
6	H	1068	0	1040	93	0
7	I	997	0	955	71	0
8	J	532	0	542	78	0
9	K	919	0	929	62	0
10	L	364	0	388	47	0
11	A	2	0	0	0	0
11	B	1	0	0	0	0
11	C	1	0	0	0	0
11	I	2	0	0	1	0
11	J	1	0	0	0	0
11	L	1	0	0	0	0
12	A	1	0	0	0	0
All	All	28161	0	28229	2150	1

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 38.

The worst 5 of 2150 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)
1:A:1364:ASN:ND2	1:A:1366:ARG:HG2	1.59	1.17
7:I:111:THR:HG22	7:I:113:ASP:H	1.05	1.17
10:L:60:ARG:HG3	10:L:61:THR:H	1.04	1.12
2:B:846:ILE:HG23	2:B:974:PRO:HG2	1.32	1.11
1:A:1161:THR:HG22	1:A:1163:ILE:H	1.06	1.10

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:106:ASP:OD1	2:B:106:ASP:OD1[2_655]	2.08	0.12

## 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%)	1138 (81%)	203 (14%)	65 (5%)	3	17
2	B	1061/1224 (87%)	868 (82%)	128 (12%)	65 (6%)	2	11
3	C	264/318 (83%)	210 (80%)	35 (13%)	19 (7%)	1	7
4	E	213/215 (99%)	184 (86%)	27 (13%)	2 (1%)	21	61
5	F	82/155 (53%)	63 (77%)	15 (18%)	4 (5%)	3	16
6	H	129/146 (88%)	91 (70%)	22 (17%)	16 (12%)	0	1
7	I	120/122 (98%)	97 (81%)	17 (14%)	6 (5%)	3	16
8	J	63/70 (90%)	53 (84%)	7 (11%)	3 (5%)	3	17
9	K	112/120 (93%)	106 (95%)	5 (4%)	1 (1%)	21	61
10	L	44/70 (63%)	28 (64%)	9 (20%)	7 (16%)	0	0
All	All	3494/4173 (84%)	2838 (81%)	468 (13%)	188 (5%)	2	14

5 of 188 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	5	GLN
1	A	35	ILE
1	A	48	ALA
1	A	55	ASP
1	A	56	PRO

### 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	1234/1520 (81%)	1165 (94%)	69 (6%)	26	62
2	B	942/1061 (89%)	884 (94%)	58 (6%)	23	59
3	C	234/274 (85%)	221 (94%)	13 (6%)	26	62
4	E	197/197 (100%)	193 (98%)	4 (2%)	63	86
5	F	74/137 (54%)	67 (90%)	7 (10%)	11	38
6	H	117/128 (91%)	108 (92%)	9 (8%)	16	50
7	I	116/116 (100%)	107 (92%)	9 (8%)	16	49
8	J	60/65 (92%)	54 (90%)	6 (10%)	9	34
9	K	99/102 (97%)	90 (91%)	9 (9%)	12	40
10	L	40/57 (70%)	35 (88%)	5 (12%)	6	22
All	All	3113/3657 (85%)	2924 (94%)	189 (6%)	23	59

5 of 189 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	480	SER
2	B	780	VAL
8	J	55	ASP
2	B	487	THR
2	B	602	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 81 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	465	ASN
2	B	744	HIS
6	H	134	ASN
2	B	513	GLN
2	B	538	ASN



### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

### 5.5 Carbohydrates [i](#)

There are no carbohydrates 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](#)

EDS was not executed - this section will therefore be empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section will therefore be empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section will therefore be empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section will therefore be empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section will therefore be empty.