wwPDB X-ray Structure Validation Summary Report (i)

Jan 30, 2024 – 12:45 PM EST

PDB ID	:	1I3Q
Title	:	RNA POLYMERASE II CRYSTAL FORM I AT 3.1 A 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

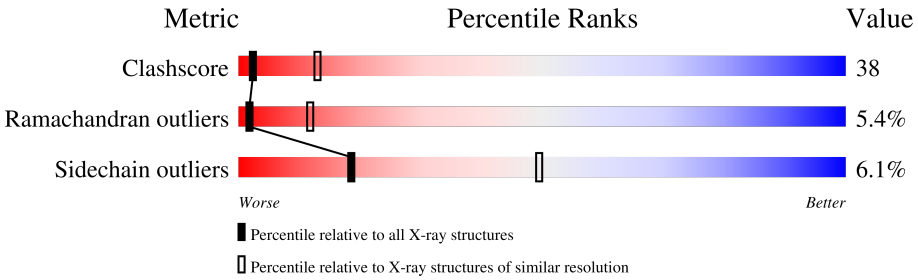
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.10 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	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)

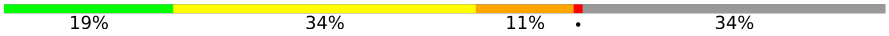
The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	1733	36% 40% 6% 18%
2	B	1224	37% 45% 6% 12%
3	C	318	34% 42% 8% 16%
4	E	215	52% 46% •
5	F	155	21% 29% 5% 46%
6	H	146	34% 46% 9% • 9%
7	I	122	45% 46% 9%
8	J	70	29% 56% 6% • 7%

Continued on next page...

Continued from previous page...

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

2 Entry composition (i)

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 SUBUNIT.

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

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

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

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

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

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II 27KD POLYPEPTIDE.

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

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

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

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

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	A	2	Total	Zn	0	0
			2	2		
11	B	1	Total	Zn	0	0
			1	1		
11	C	1	Total	Zn	0	0
			1	1		
11	I	2	Total	Zn	0	0
			2	2		

Continued on next page...

Continued from previous page...

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

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

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

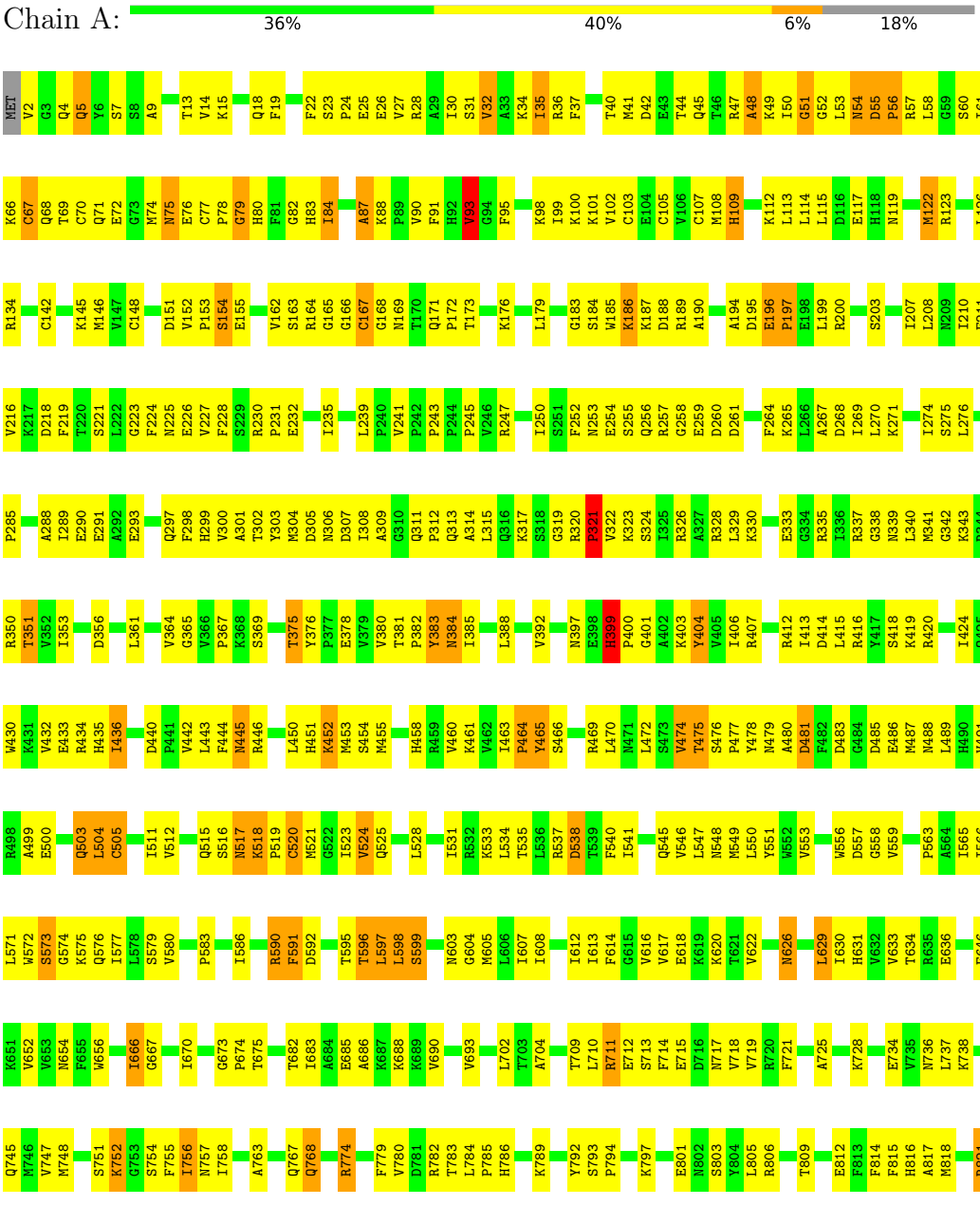
3 Residue-property plots (i)

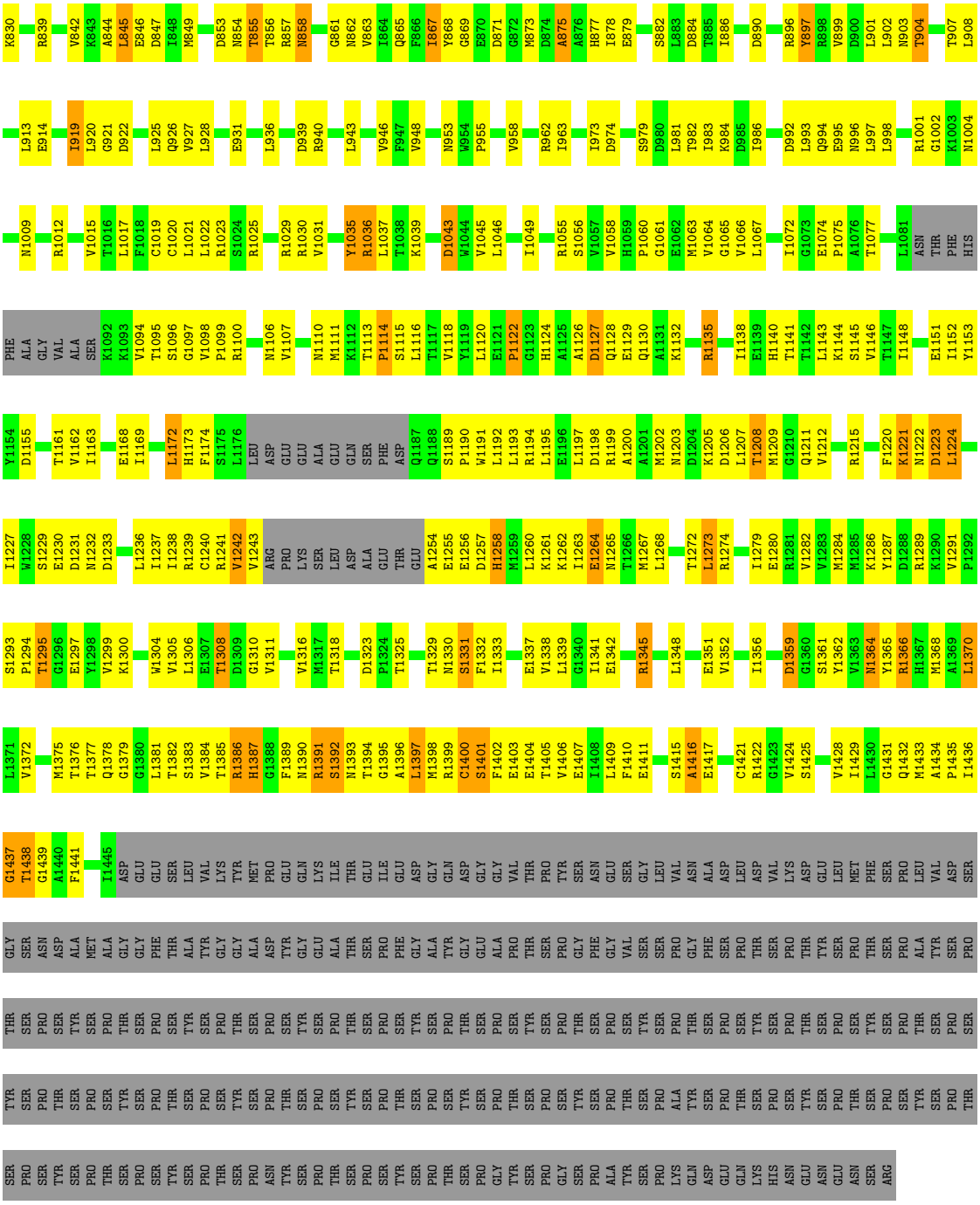
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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

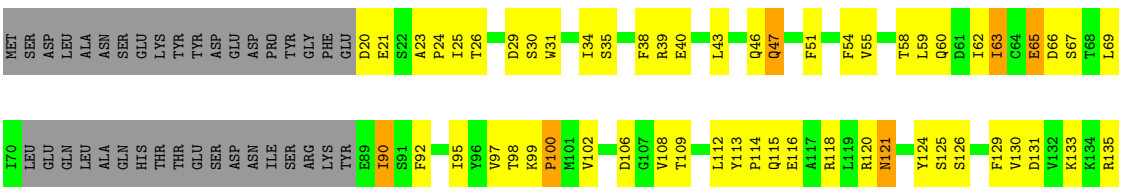
Chain A:

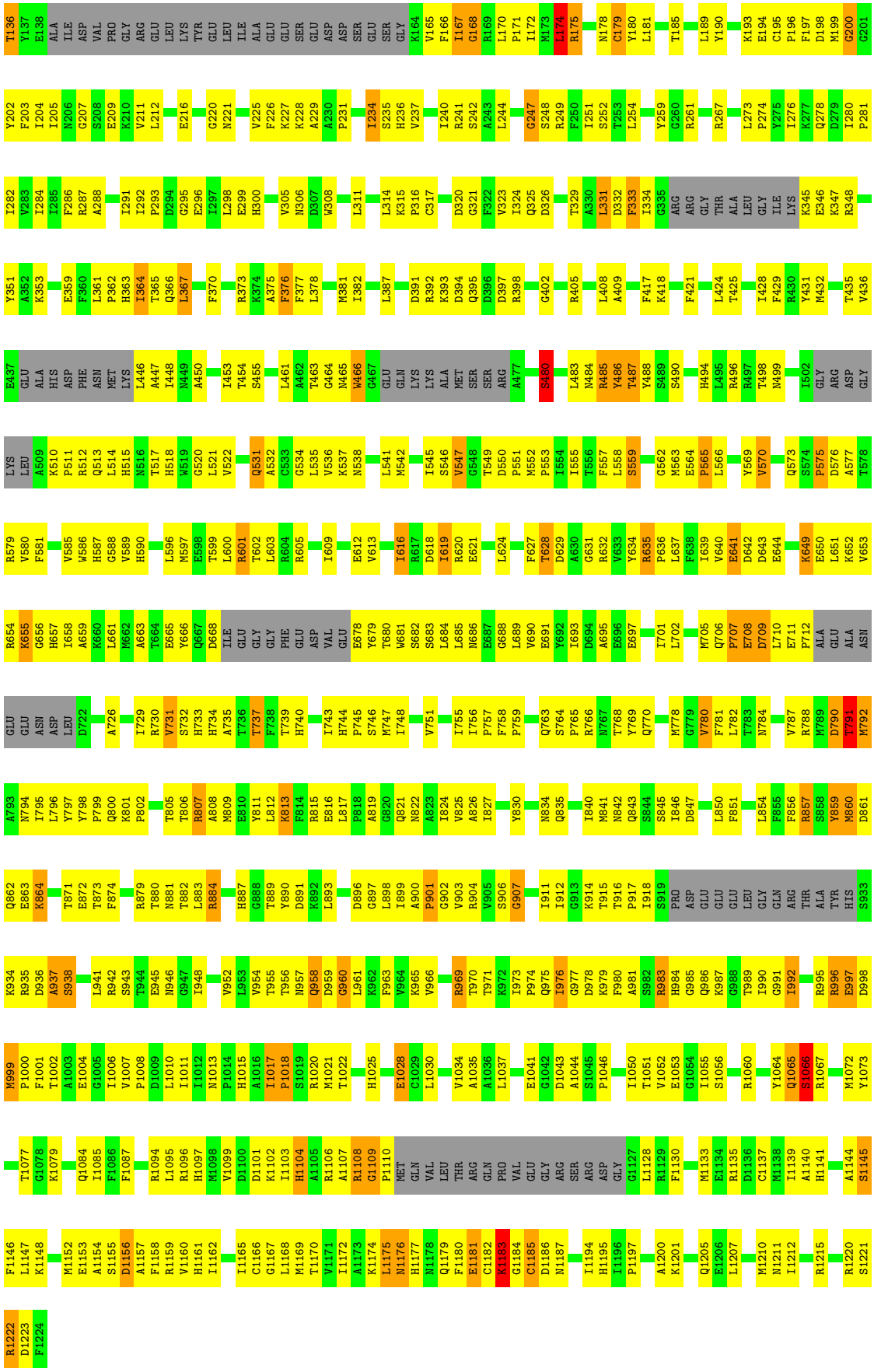




• Molecule 2: DNA-DIRECTED RNA POLYMERASE II 140KD POLYPEPTIDE

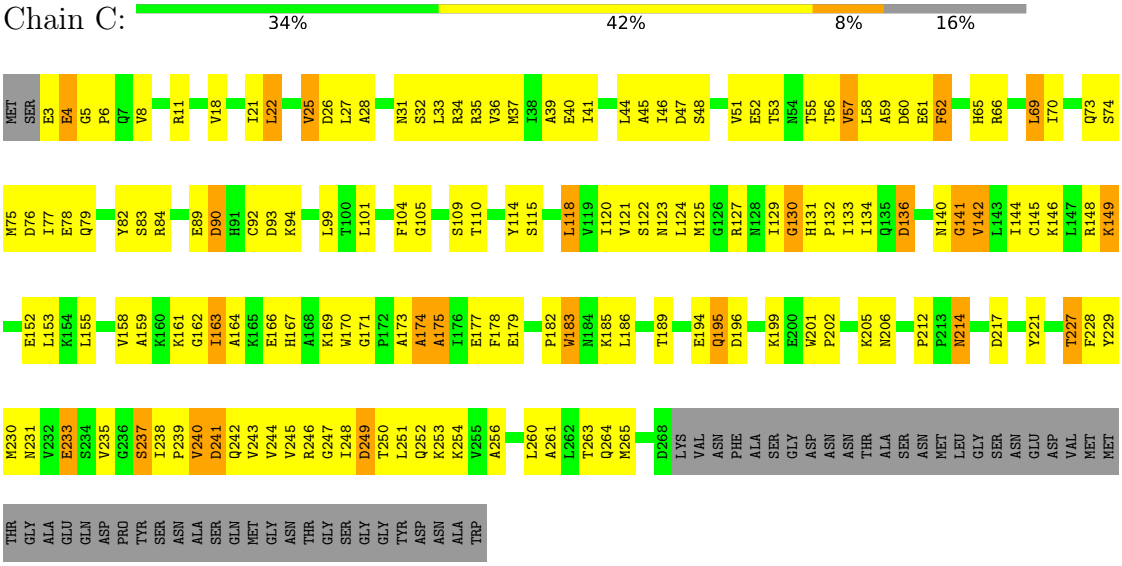
Chain B:





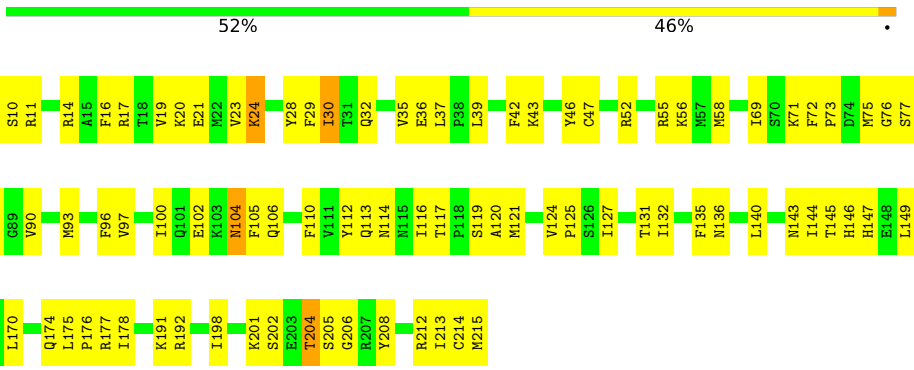
• Molecule 3: DNA-DIRECTED RNA POLYMERASE II 45KD POLYPEPTIDE

Chain C:



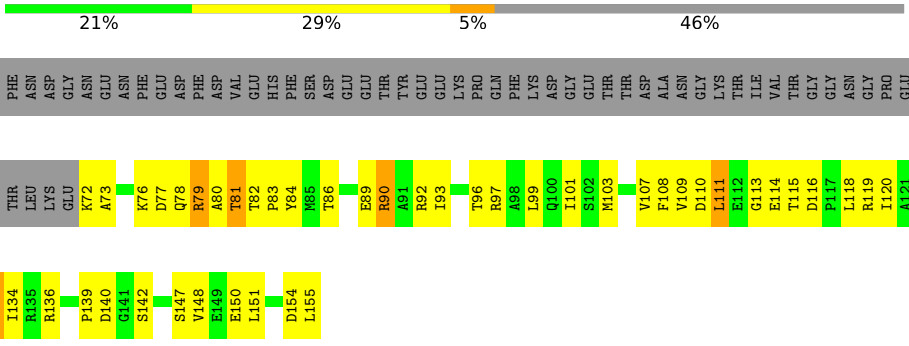
• Molecule 4: DNA-DIRECTED RNA POLYMERASE II 27KD POLYPEPTIDE

Chain E:



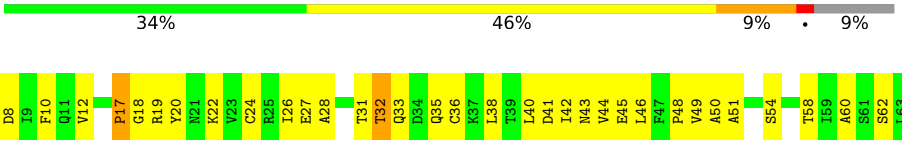
• Molecule 5: DNA-DIRECTED RNA POLYMERASE II 23KD POLYPEPTIDE

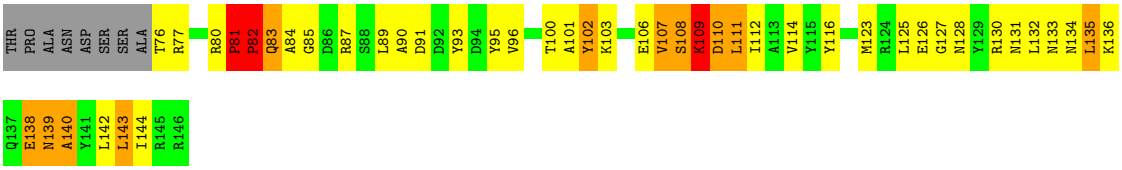
Chain F:



• Molecule 6: DNA-DIRECTED RNA POLYMERASE II 14.5KD POLYPEPTIDE

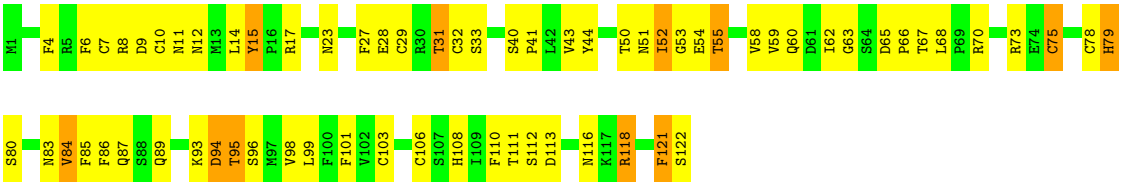
Chain H:





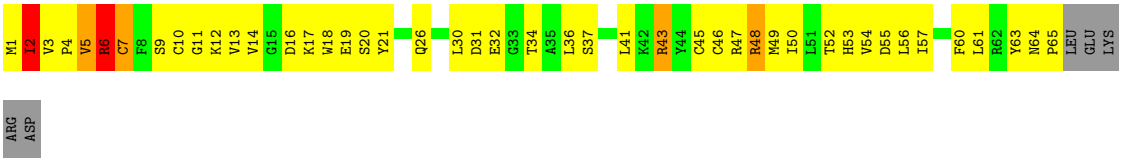
• Molecule 7: DNA-DIRECTED RNA POLYMERASE II 14.2KD POLYPEPTIDE

Chain I:



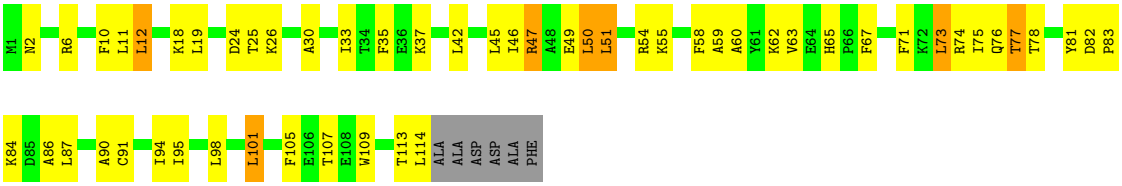
• Molecule 8: DNA-DIRECTED RNA POLYMERASE II 8.3KD POLYPEPTIDE

Chain J:



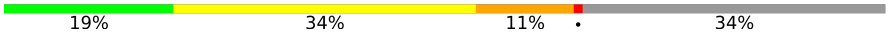
• Molecule 9: DNA-DIRECTED RNA POLYMERASE II 13.6KD POLYPEPTIDE

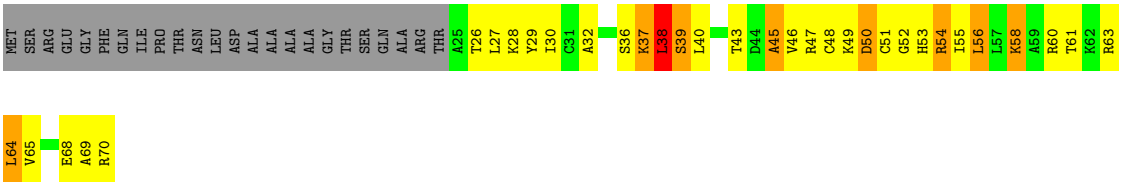
Chain K:



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

Chain L:





4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	130.70Å 224.80Å 369.40Å 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^2)	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(^{o})	Ideal(^{o})
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 (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	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%)	2 15
2	B	1061/1224 (87%)	868 (82%)	128 (12%)	65 (6%)	1 9
3	C	264/318 (83%)	210 (80%)	35 (13%)	19 (7%)	1 6
4	E	213/215 (99%)	184 (86%)	27 (13%)	2 (1%)	17 52
5	F	82/155 (53%)	63 (77%)	15 (18%)	4 (5%)	2 14
6	H	129/146 (88%)	91 (70%)	22 (17%)	16 (12%)	0 1
7	I	120/122 (98%)	97 (81%)	17 (14%)	6 (5%)	2 13
8	J	63/70 (90%)	53 (84%)	7 (11%)	3 (5%)	2 14
9	K	112/120 (93%)	106 (95%)	5 (4%)	1 (1%)	17 52
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 12

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 (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	1234/1520 (81%)	1165 (94%)	69 (6%)	21 52
2	B	942/1061 (89%)	884 (94%)	58 (6%)	18 49
3	C	234/274 (85%)	221 (94%)	13 (6%)	21 52
4	E	197/197 (100%)	193 (98%)	4 (2%)	55 80
5	F	74/137 (54%)	67 (90%)	7 (10%)	8 31
6	H	117/128 (91%)	108 (92%)	9 (8%)	13 41
7	I	116/116 (100%)	107 (92%)	9 (8%)	12 40
8	J	60/65 (92%)	54 (90%)	6 (10%)	7 28
9	K	99/102 (97%)	90 (91%)	9 (9%)	9 33
10	L	40/57 (70%)	35 (88%)	5 (12%)	4 18
All	All	3113/3657 (85%)	2924 (94%)	189 (6%)	18 49

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

Mol	Chain	Res	Type
2	B	976	ILE
4	E	204	THR
2	B	1028	GLU
3	C	57	VAL
5	F	140	ASP

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

Mol	Chain	Res	Type
3	C	102	GLN
6	H	11	GLN
3	C	123	ASN
4	E	101	GLN
7	I	12	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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 monosaccharides 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

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.