wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 09:17 AM GMT

PDB ID	:	3HOZ
Title	:	Complete RNA polymerase II elongation complex IV with a T-U mismatch and a frayed RNA 3'-guanine
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
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

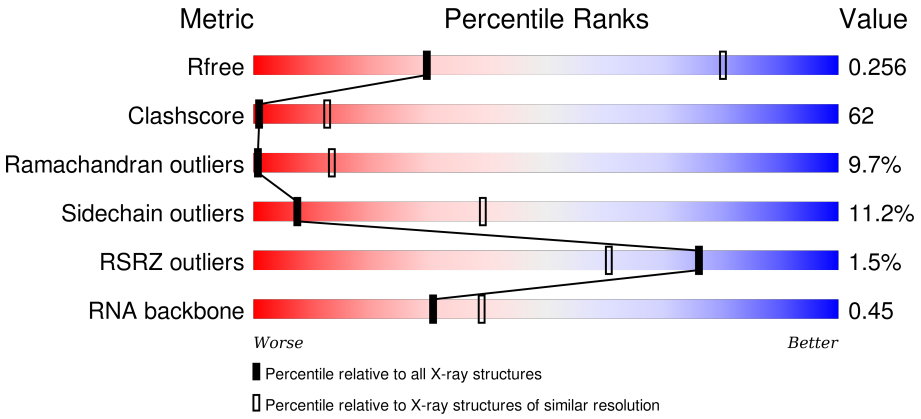
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 3.65 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.



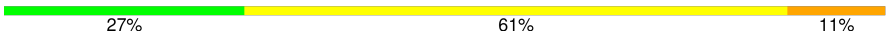
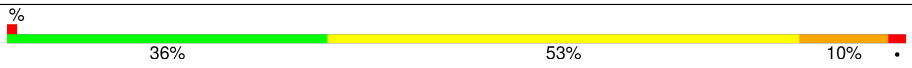
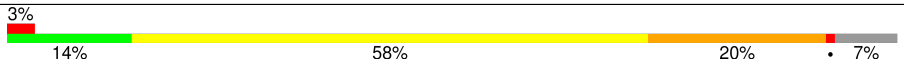
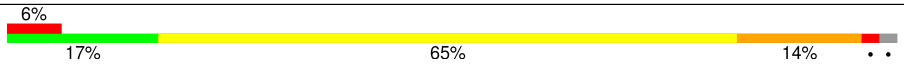
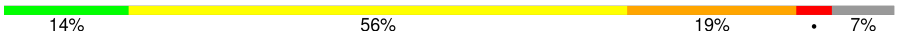
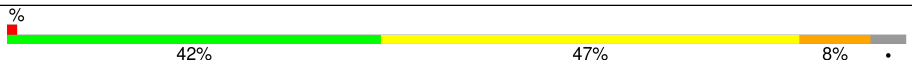
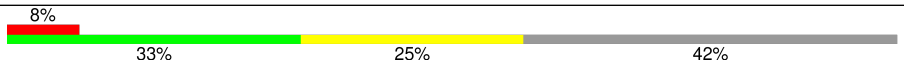
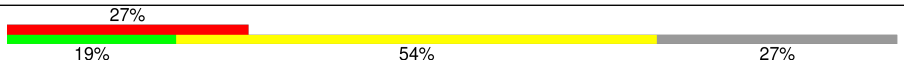
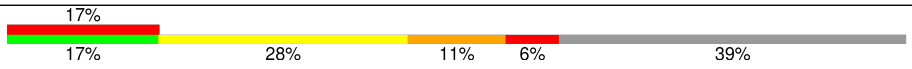
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	1010 (3.82-3.50)
Clashscore	102246	1125 (3.82-3.50)
Ramachandran outliers	100387	1079 (3.82-3.50)
Sidechain outliers	100360	1078 (3.82-3.50)
RSRZ outliers	91569	1017 (3.82-3.50)
RNA backbone	2183	1066 (4.52-2.80)

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 >=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 <=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	% 23% 47% 10% • 18%
2	B	1224	% 23% 52% 14% • 9%
3	C	347	18% 47% 11% • 23%
4	D	221	23% 45% 12% • 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

There are 17 unique types of molecules in this entry. The entry contains 31961 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	1418	Total	C	N	O	S	0	0	0
			11158	7030	1951	2115	62			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1109	Total	C	N	O	S	0	0	0
			8821	5584	1546	1636	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			

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
4	D	179	Total	C	N	O	S	0	0	0
			1443	892	258	291	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	136	Total	C	N	O	S	0	0	0
			1092	688	184	215	5			

- 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	115	Total	C	N	O	S	0	0	0
			924	593	157	172	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*CP*TP*AP*CP*TP*TP*GP*AP*GP*CP*T)-3'.

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

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
14	T	19	Total	Br	C	N	O	P	0	0	0
			387	1	185	69	114	18			

- 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*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
15	P	11	Total	C	N	O	P	0	0	0
			232	105	44	73	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).

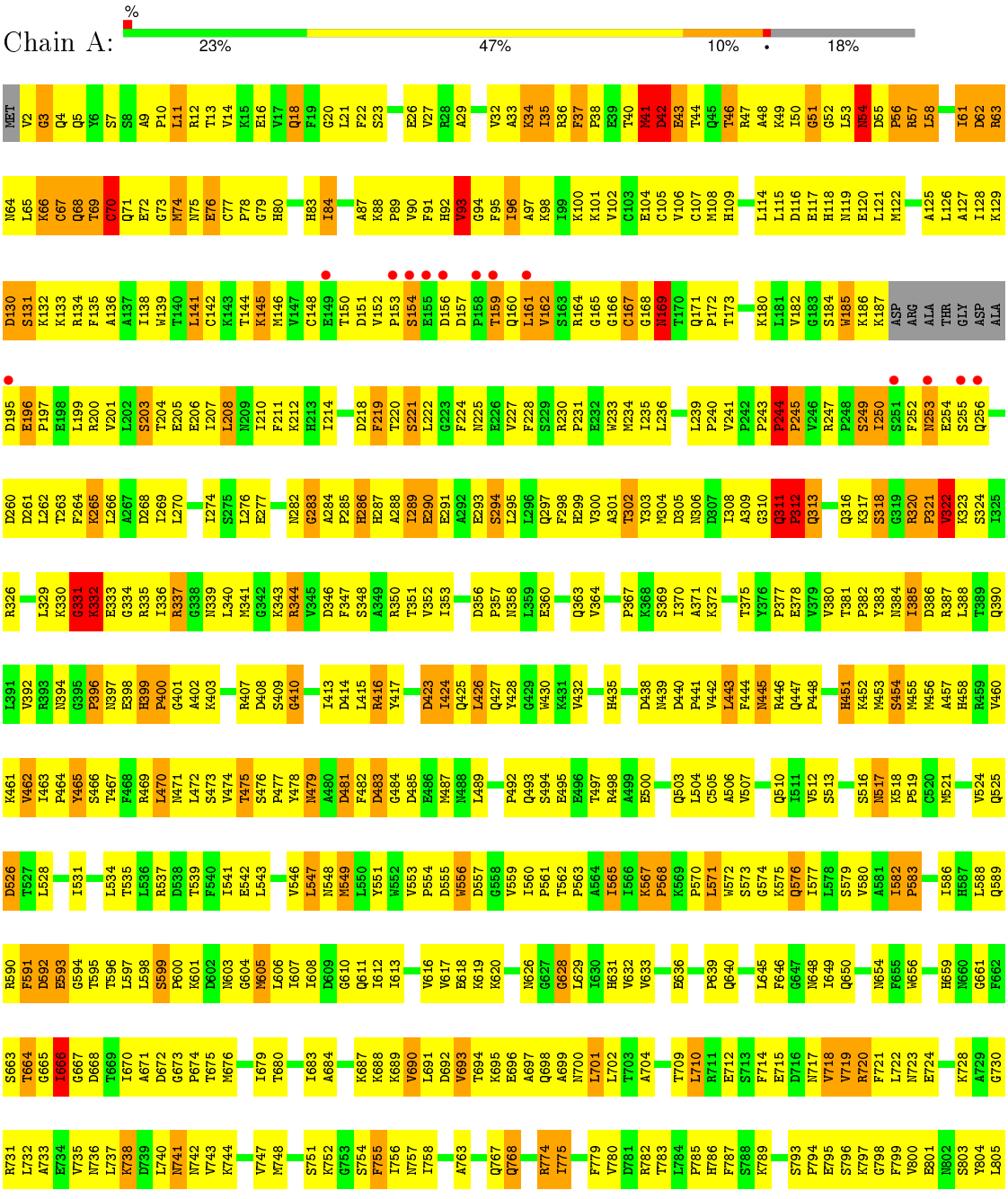
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
17	P	1	Total	Mg	0	0
			1	1		

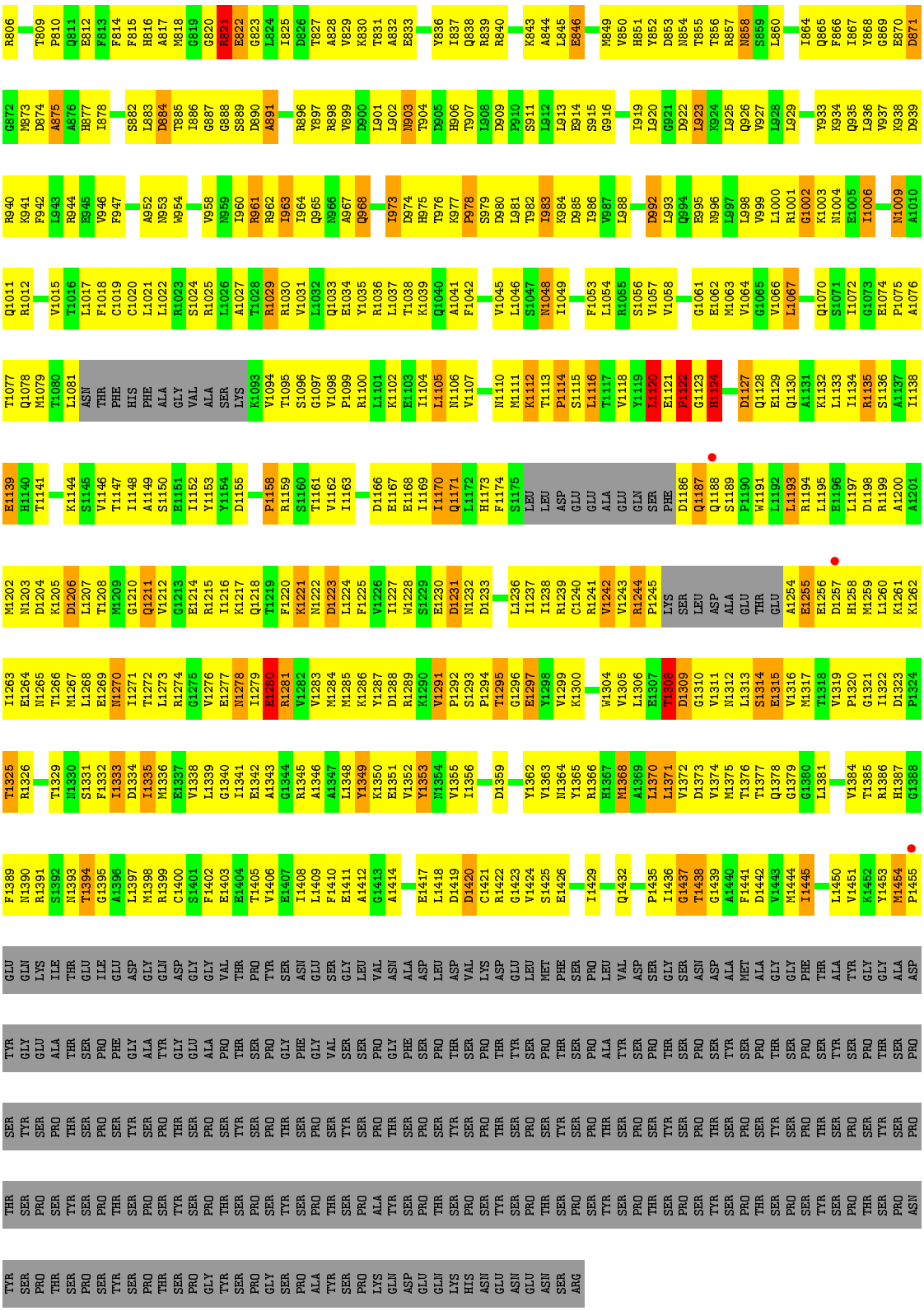
3 Residue-property plots (i)

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

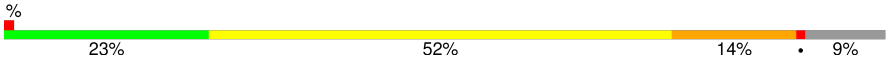
Chain A:

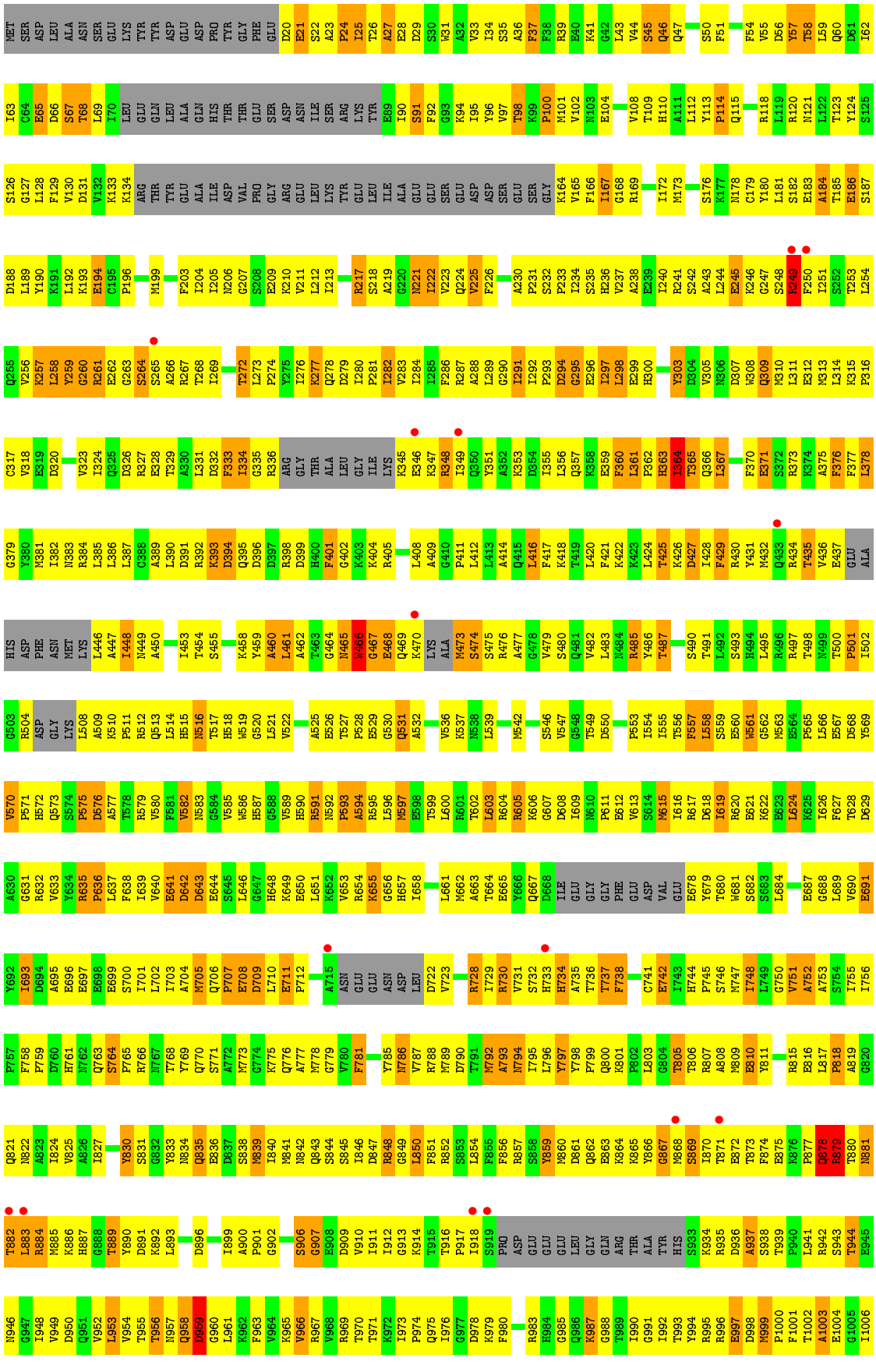


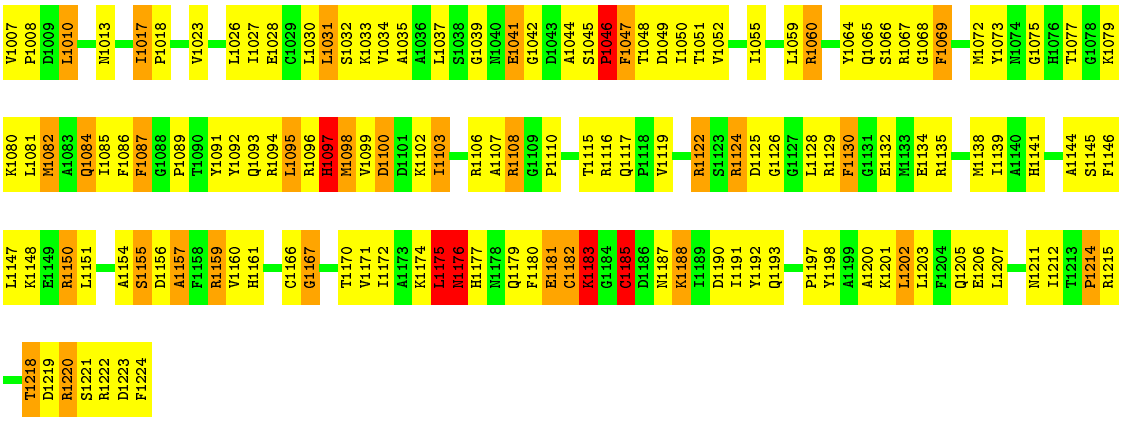


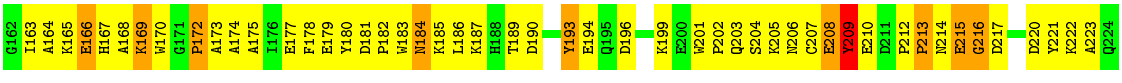
• Molecule 2: DNA-directed RNA polymerase II subunit RPB2

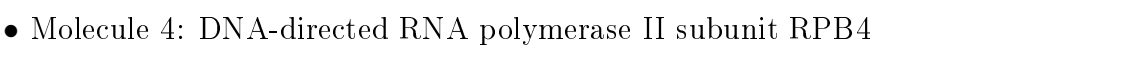
Chain B:

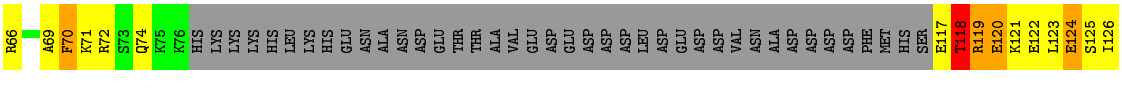






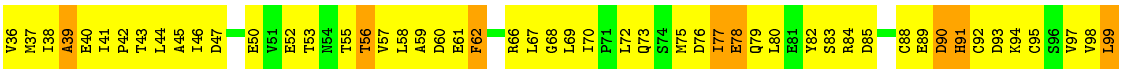


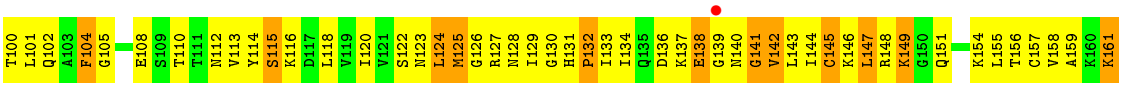


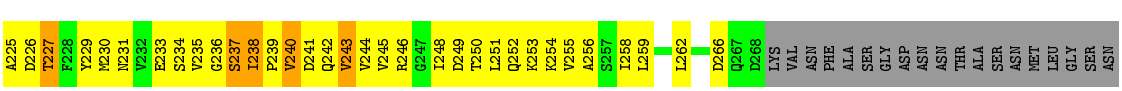


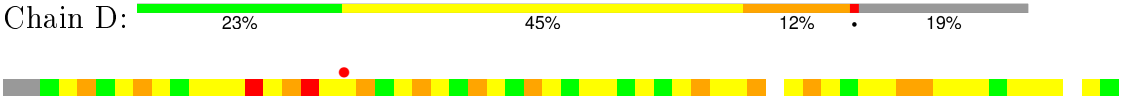
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3

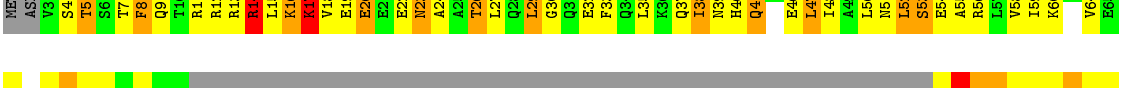
Chain C:

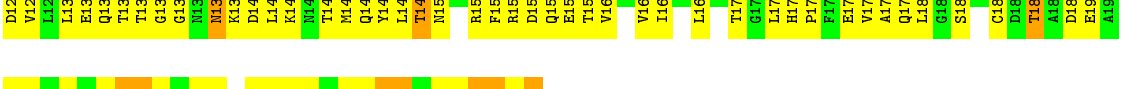






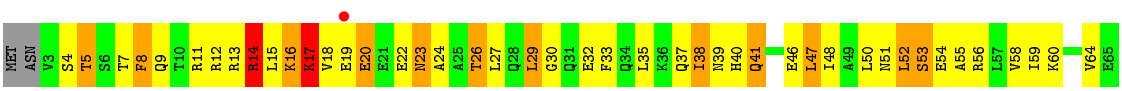


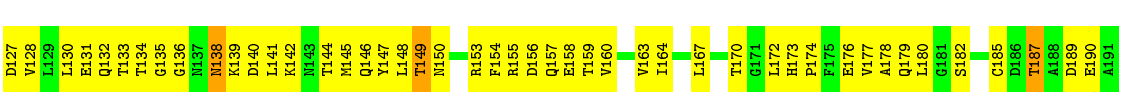




• Molecule 4: DNA-directed RNA polymerase II subunit RPB4

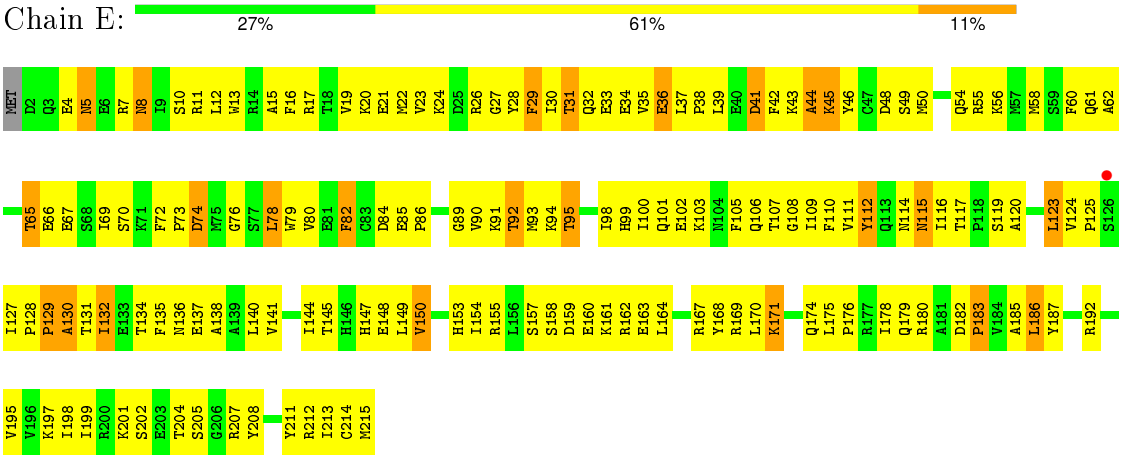
Chain D:





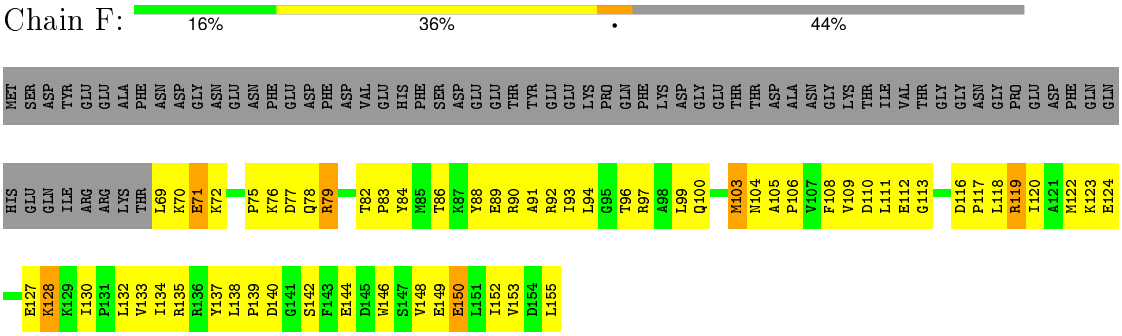
• Molecule 5: DNA-directed RNA polymerases I, II, and III subunit RPABC1

Chain E:



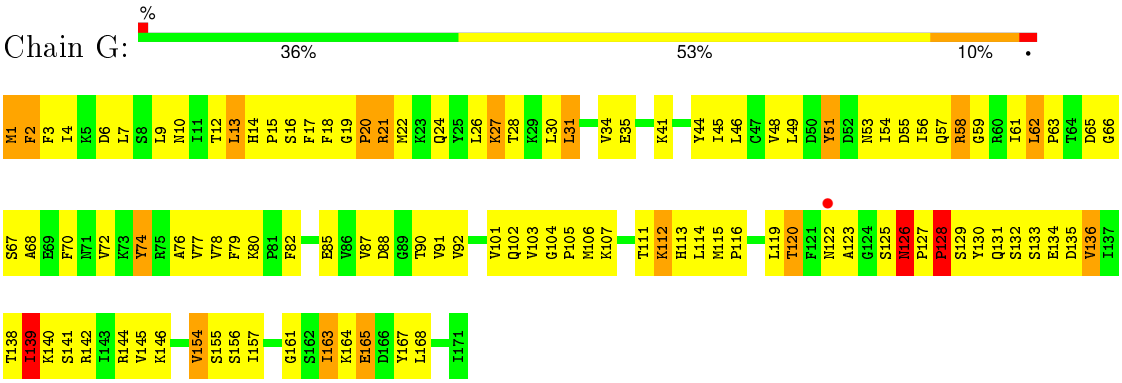
• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2

Chain F:



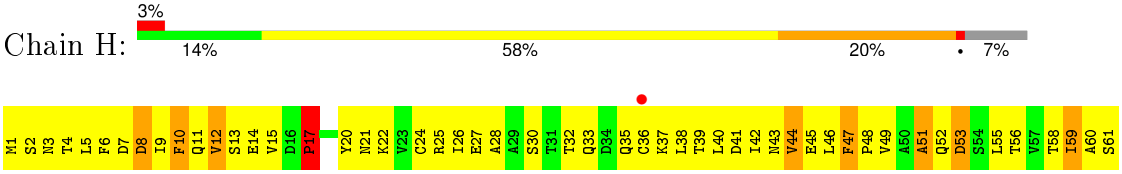
• Molecule 7: DNA-directed RNA polymerase II subunit RPB7

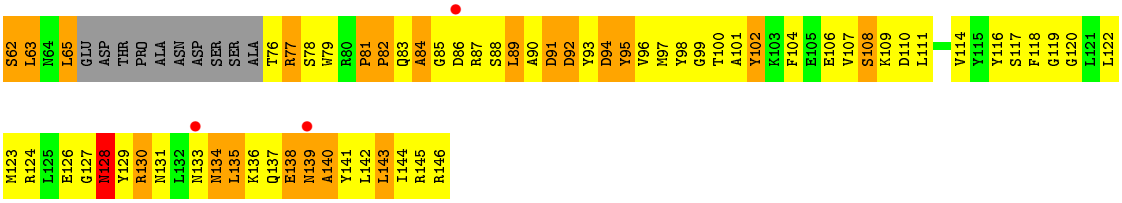
Chain G:



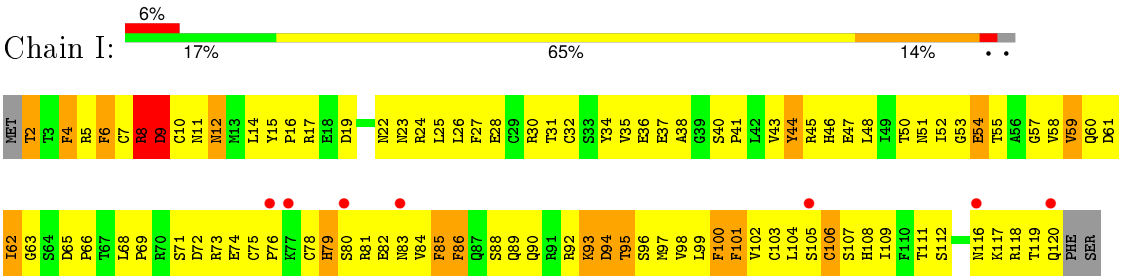
• Molecule 8: DNA-directed RNA polymerases I, II, and III subunit RPABC3

Chain H:

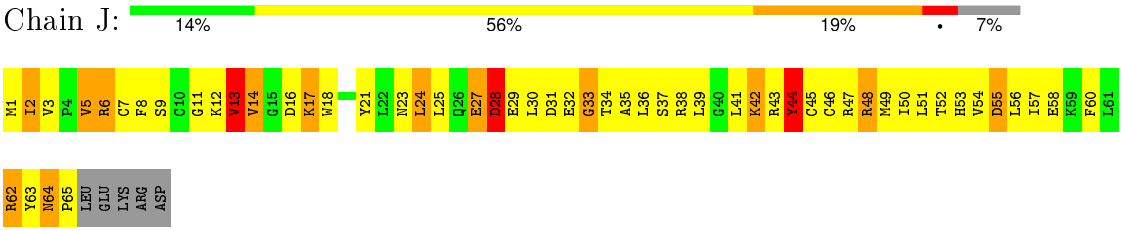




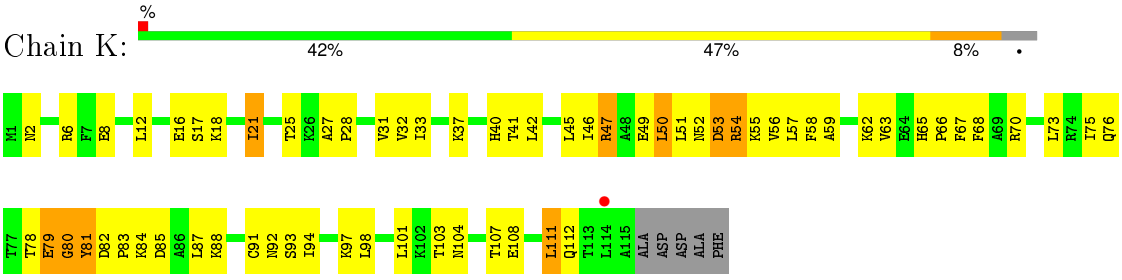
• Molecule 9: DNA-directed RNA polymerase II subunit RPB9



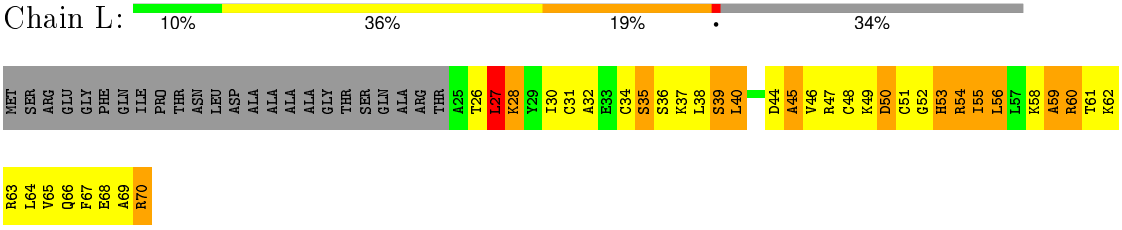
• Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5



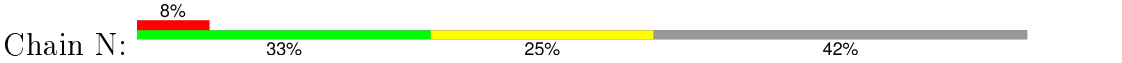
• Molecule 11: DNA-directed RNA polymerase II subunit RPB11



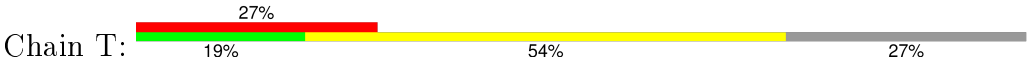
• Molecule 12: DNA-directed RNA polymerases I, II, and III subunit RPABC4

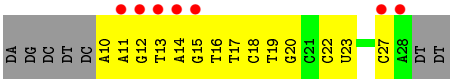


• Molecule 13: 5'-D(*AP*CP*TP*AP*CP*TP*TP*GP*AP*GP*CP*T)-3'



• Molecule 14: 5'-D(*AP*GP*CP*TP*C*AP*AP*GP*TP*AP*GP*TP*TP*CP*TP*GP*CP*CP*(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*G)-3'

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	221.43Å 393.75Å 281.82Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.65 49.84 – 3.65	Depositor EDS
% Data completeness (in resolution range)	99.9 (50.00-3.65) 100.0 (49.84-3.65)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	3.21 (at 3.67Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.210 , 0.253 0.213 , 0.256	Depositor DCC
R_{free} test set	2674 reflections (2.01%)	DCC
Wilson B-factor (Å^2)	88.8	Xtriage
Anisotropy	0.057	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 112.9	EDS
Estimated twinning fraction	0.017 for 1/2*h-1/2*k,-3/2*h-1/2*k,-l 0.025 for 1/2*h+1/2*k,3/2*h-1/2*k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Outliers	0 of 135971 reflections	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	31961	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.01% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	0.51	0/11358	0.79	4/15360 (0.0%)
2	B	0.49	0/8991	0.74	4/12121 (0.0%)
3	C	0.50	0/2133	0.74	1/2891 (0.0%)
4	D	0.48	0/1453	0.77	1/1947 (0.1%)
5	E	0.48	0/1788	0.71	2/2406 (0.1%)
6	F	0.57	0/717	0.83	1/967 (0.1%)
7	G	0.54	0/1368	0.81	1/1844 (0.1%)
8	H	0.45	0/1110	0.74	0/1502
9	I	0.44	0/989	0.72	0/1331
10	J	0.51	0/541	0.85	1/727 (0.1%)
11	K	0.49	0/942	0.68	0/1272
12	L	0.56	0/365	0.82	0/485
13	N	0.60	0/152	0.90	0/232
14	T	0.58	0/410	0.82	0/629
15	P	0.57	0/259	0.82	1/402 (0.2%)
All	All	0.50	0/32576	0.76	16/44116 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
10	J	0	1

There are no bond length outliers.

The worst 5 of 16 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	D	26	THR	N-CA-C	-6.48	93.51	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	C	39	ALA	N-CA-C	6.32	128.07	111.00
1	A	331	GLY	N-CA-C	5.96	128.00	113.10
7	G	65	ASP	N-CA-C	-5.92	95.02	111.00
1	A	3	GLY	N-CA-C	-5.78	98.65	113.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
10	J	44	TYR	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	11158	0	11228	1381	0
2	B	8821	0	8850	1234	0
3	C	2095	0	2051	306	0
4	D	1443	0	1466	213	0
5	E	1752	0	1776	214	0
6	F	705	0	731	92	0
7	G	1340	0	1357	168	0
8	H	1092	0	1069	179	0
9	I	971	0	929	137	0
10	J	532	0	542	112	0
11	K	924	0	934	105	0
12	L	363	0	388	83	0
13	N	137	0	82	4	0
14	T	387	0	214	25	0
15	P	232	0	122	14	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	P	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	31961	0	31739	3920	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 62.

The worst 5 of 3920 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:508:LEU:HD13	2:B:510:LYS:HE2	1.26	1.16
1:A:53:LEU:HD23	1:A:54:ASN:N	1.61	1.16
2:B:744:HIS:HD2	2:B:745:PRO:HD2	1.07	1.14
1:A:1208:THR:HB	1:A:1211:GLN:HG3	1.24	1.13
2:B:559:SER:HA	2:B:563:MET:HB3	1.15	1.13

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	1408/1733 (81%)	1012 (72%)	262 (19%)	134 (10%)	1 14
2	B	1089/1224 (89%)	779 (72%)	201 (18%)	109 (10%)	1 12
3	C	264/347 (76%)	186 (70%)	51 (19%)	27 (10%)	1 12
4	D	175/221 (79%)	121 (69%)	39 (22%)	15 (9%)	1 16
5	E	212/215 (99%)	154 (73%)	42 (20%)	16 (8%)	1 20
6	F	85/155 (55%)	69 (81%)	14 (16%)	2 (2%)	7 51
7	G	169/171 (99%)	145 (86%)	13 (8%)	11 (6%)	1 25
8	H	132/146 (90%)	85 (64%)	23 (17%)	24 (18%)	0 3
9	I	117/122 (96%)	79 (68%)	29 (25%)	9 (8%)	1 19

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
10	J	63/70 (90%)	39 (62%)	11 (18%)	13 (21%)	0 2
11	K	113/120 (94%)	87 (77%)	22 (20%)	4 (4%)	4 43
12	L	44/70 (63%)	23 (52%)	9 (20%)	12 (27%)	0 0
All	All	3871/4594 (84%)	2779 (72%)	716 (18%)	376 (10%)	1 13

5 of 376 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	4	GLN
1	A	41	MET
1	A	43	GLU
1	A	48	ALA
1	A	54	ASN

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	1241/1520 (82%)	1109 (89%)	132 (11%)	8 42
2	B	963/1061 (91%)	841 (87%)	122 (13%)	5 32
3	C	234/299 (78%)	206 (88%)	28 (12%)	6 35
4	D	161/200 (80%)	139 (86%)	22 (14%)	4 30
5	E	196/197 (100%)	180 (92%)	16 (8%)	14 54
6	F	77/137 (56%)	73 (95%)	4 (5%)	29 70
7	G	152/152 (100%)	137 (90%)	15 (10%)	10 45
8	H	120/128 (94%)	104 (87%)	16 (13%)	5 31
9	I	113/116 (97%)	97 (86%)	16 (14%)	4 28
10	J	60/65 (92%)	55 (92%)	5 (8%)	14 53
11	K	99/102 (97%)	92 (93%)	7 (7%)	18 60
12	L	40/57 (70%)	36 (90%)	4 (10%)	9 44
All	All	3456/4034 (86%)	3069 (89%)	387 (11%)	7 39

5 of 387 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	493	SER
2	B	909	ASP
9	I	6	PHE
2	B	557	PHE
2	B	737	THR

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

Mol	Chain	Res	Type
1	A	1432	GLN
2	B	657	HIS
7	G	131	GLN
2	B	60	GLN
2	B	366	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
15	P	10/18 (55%)	2 (20%)	1 (10%)

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
15	P	3	A
15	P	11	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
15	P	2	A

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	15,14	13,21,22	4.60	4 (30%)	16,30,33	3.98	2 (12%)

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	15,14	-	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)
14	T	23	BRU	BR-C5	-15.07	1.50	1.90
14	T	23	BRU	C6-N1	2.52	1.38	1.35
14	T	23	BRU	C4-N3	3.21	1.39	1.33
14	T	23	BRU	C4-C5	5.32	1.45	1.38

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
14	T	23	BRU	C5-C4-N3	-7.91	115.55	124.00
14	T	23	BRU	C4-N3-C2	13.47	126.89	115.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
14	T	23	BRU	2	0

5.5 Carbohydrates (i)

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^{th} 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^2)	Q<0.9
1	A	1418/1733 (81%)	-0.11	16 (1%) 82 68	23, 70, 112, 141	0
2	B	1109/1224 (90%)	-0.09	15 (1%) 78 62	23, 81, 123, 141	0
3	C	266/347 (76%)	-0.14	1 (0%) 93 88	34, 69, 105, 119	0
4	D	179/221 (80%)	-0.02	1 (0%) 90 82	37, 79, 118, 131	0
5	E	214/215 (99%)	0.13	1 (0%) 91 85	41, 97, 129, 137	0
6	F	87/155 (56%)	-0.38	0 100 100	19, 46, 77, 86	0
7	G	171/171 (100%)	-0.06	1 (0%) 90 82	48, 64, 104, 113	0
8	H	136/146 (93%)	0.48	4 (2%) 55 37	80, 106, 127, 135	0
9	I	119/122 (97%)	0.31	7 (5%) 26 14	65, 100, 125, 143	0
10	J	65/70 (92%)	-0.34	0 100 100	49, 65, 92, 105	0
11	K	115/120 (95%)	-0.08	1 (0%) 85 74	34, 73, 93, 122	0
12	L	46/70 (65%)	0.22	0 100 100	48, 108, 125, 132	0
13	N	7/12 (58%)	1.66	1 (14%) 4 2	135, 140, 151, 157	0
14	T	18/26 (69%)	1.48	7 (38%) 0 1	117, 144, 155, 155	0
15	P	11/18 (61%)	1.70	3 (27%) 1 1	125, 133, 152, 156	0
All	All	3961/4650 (85%)	-0.04	58 (1%) 76 61	19, 76, 121, 157	0

The worst 5 of 58 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1455	PRO	3.9
15	P	1	C	3.9
14	T	28	DA	3.8
2	B	733	HIS	3.7
1	A	255	SER	3.6

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^{th} 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^2)	Q<0.9
14	BRU	T	23	20/21	0.67	0.36	-	136,142,145,146	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^{th} 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^2)	Q<0.9
17	MG	P	2458	1/1	0.84	0.32	0.18	180,180,180,180	0
16	ZN	J	1066	1/1	1.00	0.20	-0.36	53,53,53,53	0
16	ZN	B	2225	1/1	0.99	0.17	-0.71	30,30,30,30	0
16	ZN	C	1269	1/1	1.00	0.10	-1.57	40,40,40,40	0
16	ZN	A	2457	1/1	1.00	0.13	-1.69	31,31,31,31	0
16	ZN	L	1071	1/1	0.98	0.05	-1.92	100,100,100,100	0
16	ZN	I	1121	1/1	0.99	0.08	-2.33	76,76,76,76	0
16	ZN	I	1122	1/1	0.95	0.14	-2.63	135,135,135,135	0
16	ZN	A	2456	1/1	0.97	0.07	-4.04	88,88,88,88	0

6.5 Other polymers (i)

There are no such residues in this entry.