wwPDB X-ray Structure Validation Summary Report (i)

Aug 25, 2026 – 01:37 am BST

PDB ID	:	2Y0S / pdb_00002y0s
Title	:	Crystal structure of Sulfolobus shibatae RNA polymerase in P21 space group
Authors	:	Wojtas, M.; Peralta, B.; Ondiviela, M.; Mogni, M.; Bell, S.D.; Abrescia, N.G.A.
Deposited on	:	2010-12-07
Resolution	:	3.80 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:

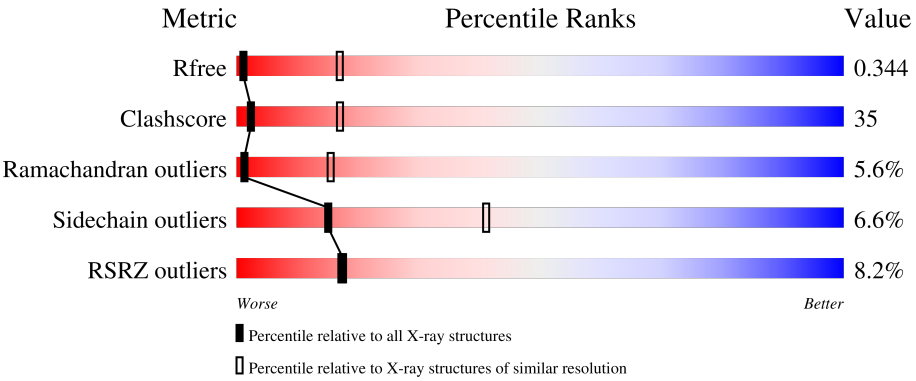
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.015 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.50

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 3.80 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}	180053	1065 (3.96-3.64)
Clashscore	190562	1012 (3.94-3.66)
Ramachandran outliers	187476	1048 (3.96-3.64)
Sidechain outliers	187428	1043 (3.96-3.64)
RSRZ outliers	180081	1064 (3.96-3.64)

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% 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	880	9% 45% 41% 8% • 6%
1	W	880	7% 43% 43% 7% • 6%
2	B	1131	8% 41% 45% 10% • •
2	R	1131	8% 41% 46% 8% • •
3	C	395	11% 39% 42% 12% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	Y	395	6% 39% 42% 12% 7%
4	D	265	7% 50% 45%
4	S	265	5% 53% 42%
5	E	180	9% 37% 44% 8% 9%
5	T	180	10% 38% 44% 7% 9%
6	F	113	4% 50% 26% 20%
6	U	113	3% 50% 27% 20%
7	G	132	5% 39% 37% 8% 17%
7	V	132	5% 39% 39% 5% 17%
8	H	84	5% 43% 40% 5% 12%
8	Z	84	6% 44% 38% 6% 12%
9	I	95	4% 36% 41% 8% 14%
9	K	95	13% 38% 39% 7% 14%
10	J	104	5% 16% 17% 62%
10	Q	104	6% 15% 16% 6% 62%
11	L	92	4% 51% 47%
11	M	92	2% 54% 42%
12	N	66	8% 50% 39% 6%
12	O	66	5% 50% 41% 6%
13	P	48	8% 40% 42% 12% 6%
13	X	48	8% 46% 38% 10% 6%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
14	ZN	W	1881	-	-	X	-
16	F3S	D	1001	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
16	F3S	S	1001	-	-	X	-

2 Entry composition (i)

There are 16 unique types of molecules in this entry. The entry contains 52472 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	831	Total	C	N	O	S	0	0	0
			6620	4211	1170	1212	27			
1	W	831	Total	C	N	O	S	0	0	0
			6620	4211	1170	1212	27			

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1085	Total	C	N	O	S	0	0	0
			8615	5461	1526	1599	29			
2	R	1085	Total	C	N	O	S	0	0	0
			8615	5461	1526	1599	29			

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT A".

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	368	Total	C	N	O	S	0	0	0
			2845	1803	486	548	8			
3	Y	368	Total	C	N	O	S	0	0	0
			2845	1803	486	548	8			

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT D.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	262	Total	C	N	O	S	0	0	0
			2081	1338	336	394	13			
4	S	262	Total	C	N	O	S	0	0	0
			2081	1338	336	394	13			

- Molecule 5 is a protein called RNA POLYMERASE SUBUNIT 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	164	Total	C	N	O	S	0	0	0
			1297	833	219	240	5			
5	T	164	Total	C	N	O	S	0	0	0
			1297	833	219	240	5			

- Molecule 6 is a protein called RNA POLYMERASE SUBUNIT 7.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	90	Total	C	N	O	S	0	0	0
			701	439	114	145	3			
6	U	90	Total	C	N	O	S	0	0	0
			701	439	114	145	3			

- Molecule 7 is a protein called RNA POLYMERASE SUBUNIT 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	110	Total	C	N	O	S	0	0	0
			878	558	147	169	4			
7	V	110	Total	C	N	O	S	0	0	0
			878	558	147	169	4			

- Molecule 8 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT H.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
8	H	74	Total	C	N	O	0	0	0
			609	396	108	105			
8	Z	74	Total	C	N	O	0	0	0
			609	396	108	105			

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT K.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	82	Total	C	N	O	S	0	0	0
			653	419	117	116	1			
9	K	82	Total	C	N	O	S	0	0	0
			653	419	117	116	1			

- Molecule 10 is a protein called RNA POLYMERASE SUBUNIT 13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	39	Total	C	N	O	S	0	0	0
			332	210	54	67	1			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	Q	39	Total	C	N	O	S	0	0	0
			332	210	54	67	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
J	40	ASN	ASP	conflict	UNP B8YB65
Q	40	ASN	ASP	conflict	UNP B8YB65

- Molecule 11 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	L	91	Total	C	N	O	S	0	0	0
			707	454	114	137	2			
11	M	91	Total	C	N	O	S	0	0	0
			707	454	114	137	2			

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASE SUBUNIT N.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	N	64	Total	C	N	O	S	0	0	0
			514	327	93	87	7			
12	O	64	Total	C	N	O	S	0	0	0
			514	327	93	87	7			

- Molecule 13 is a protein called RNA POLYMERASE SUBUNIT 12.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	P	45	Total	C	N	O	S	0	0	0
			369	245	63	56	5			
13	X	45	Total	C	N	O	S	0	0	0
			369	245	63	56	5			

- Molecule 14 is ZINC ION (CCD ID: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	2	Total	Zn	0	0
			2	2		
14	B	3	Total	Zn	0	0
			3	3		
14	N	1	Total	Zn	0	0
			1	1		

Continued on next page...

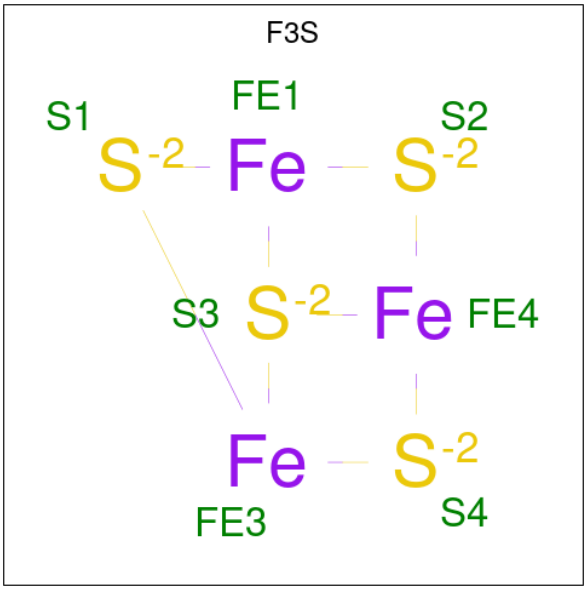
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
14	O	1	Total Zn 1 1	0	0
14	P	1	Total Zn 1 1	0	0
14	R	3	Total Zn 3 3	0	0
14	W	2	Total Zn 2 2	0	0
14	X	1	Total Zn 1 1	0	0

- Molecule 15 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	A	1	Total Mg 1 1	0	0
15	W	1	Total Mg 1 1	0	0

- Molecule 16 is FE3-S4 CLUSTER (CCD ID: F3S) (formula: Fe_3S_4).



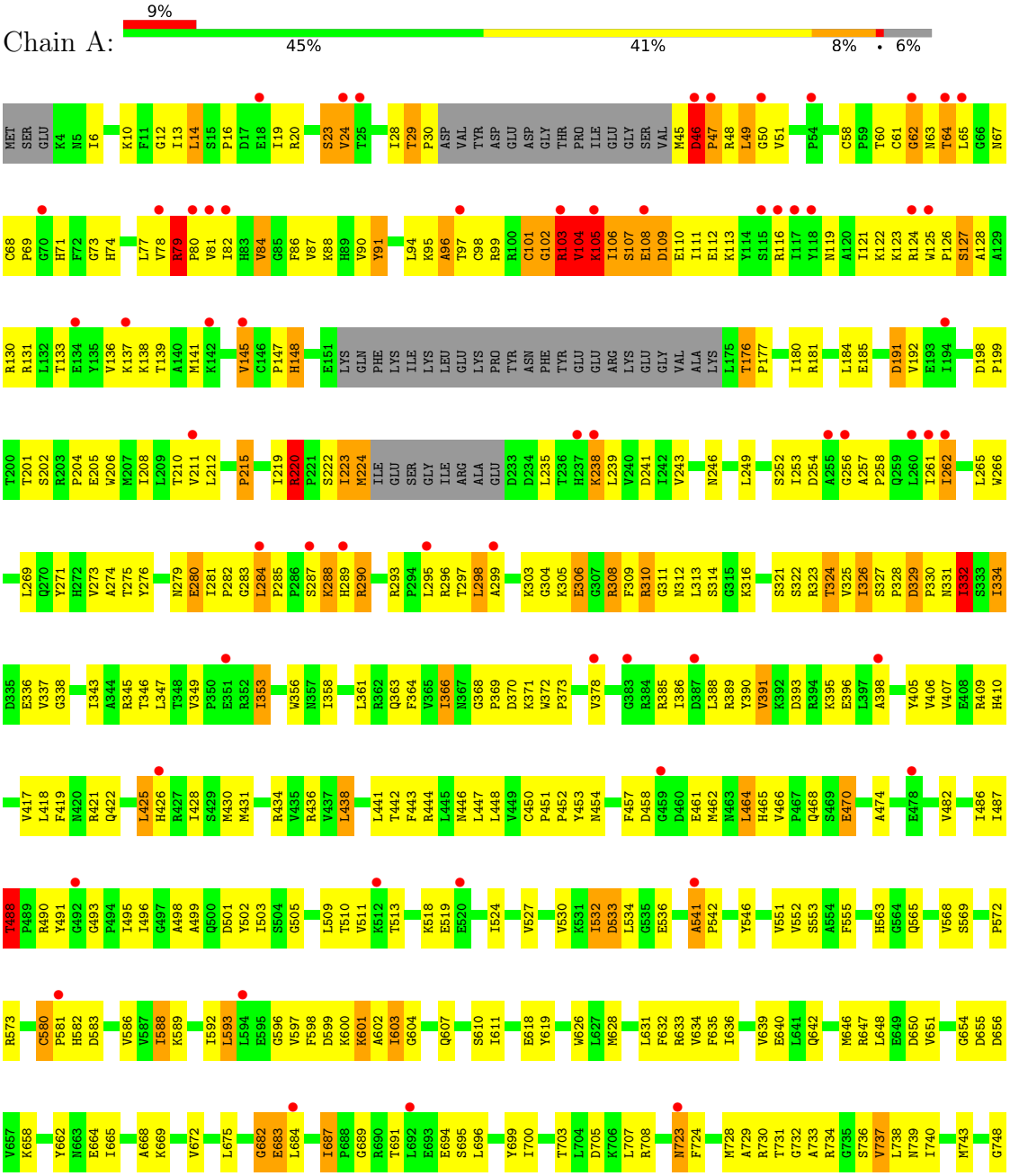
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	D	1	Total Fe S 7 3 4	0	0
16	S	1	Total Fe S 7 3 4	0	0

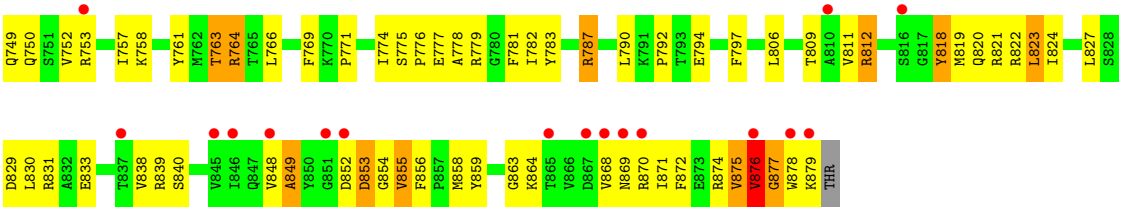
3 Residue-property plots (i)

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 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

Chain A:

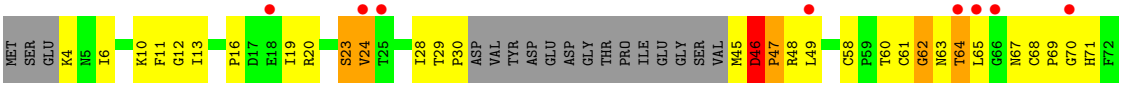


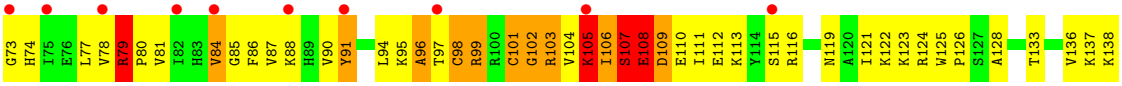


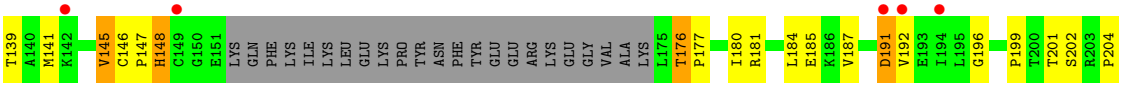
• Molecule 1: DNA-DIRECTED RNA POLYMERASE

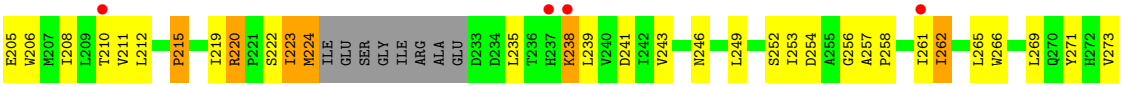
Chain W:

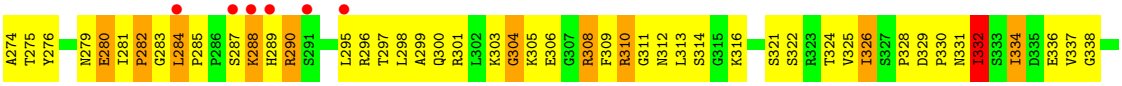


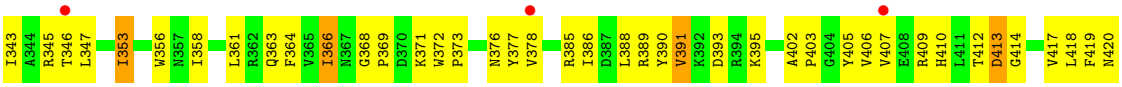




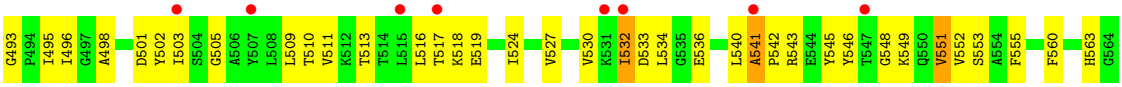


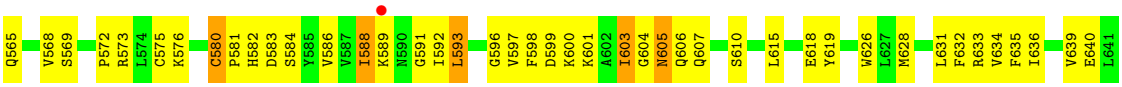


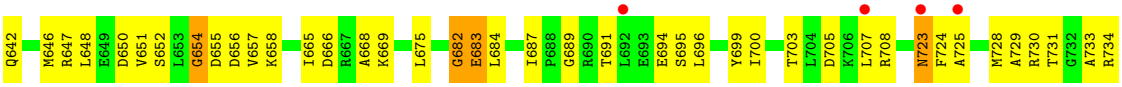


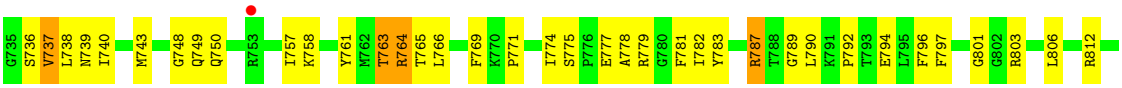




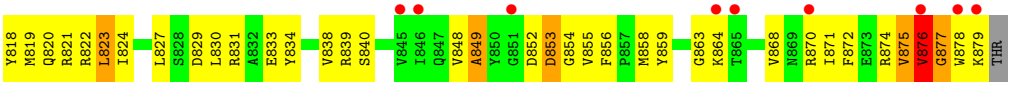








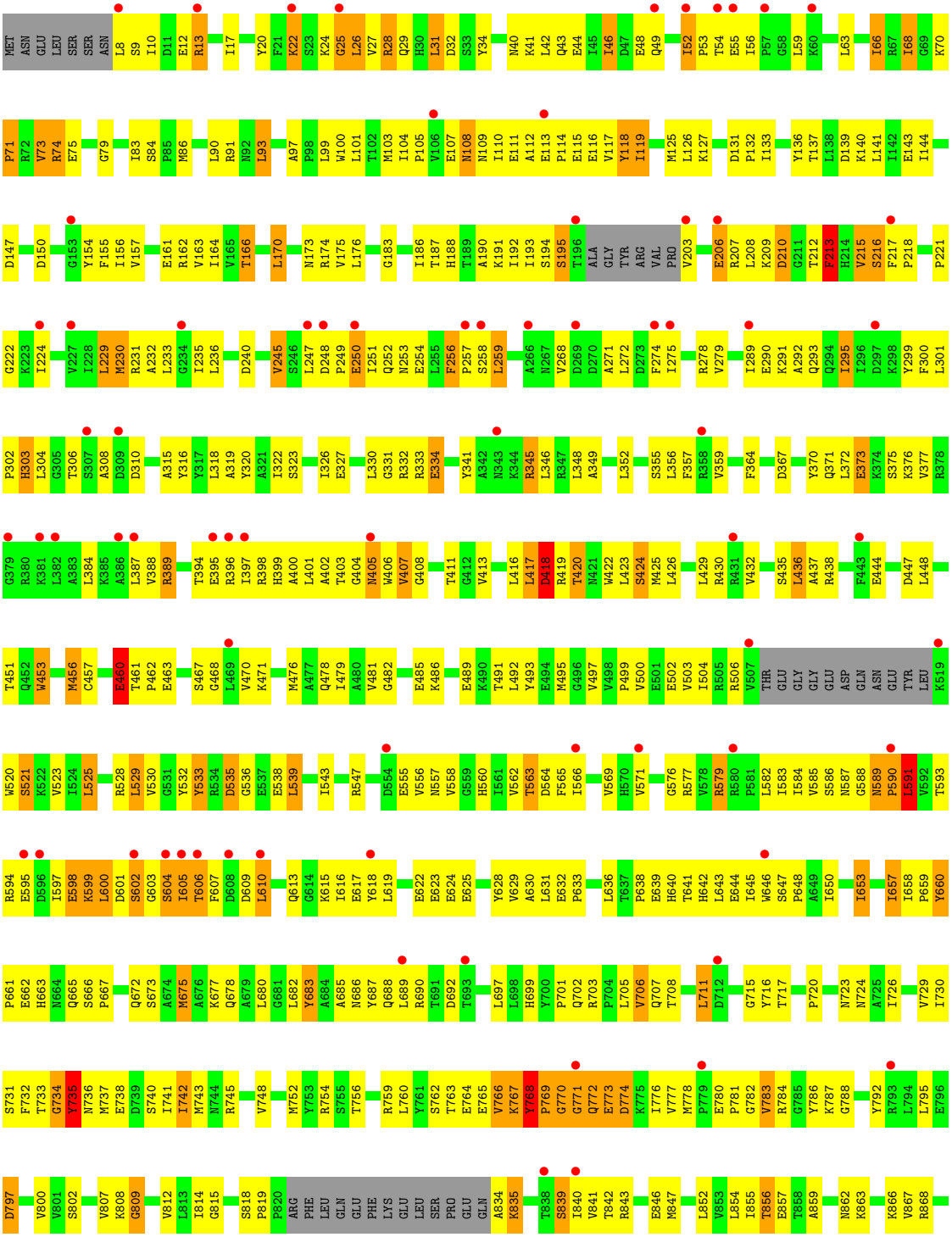


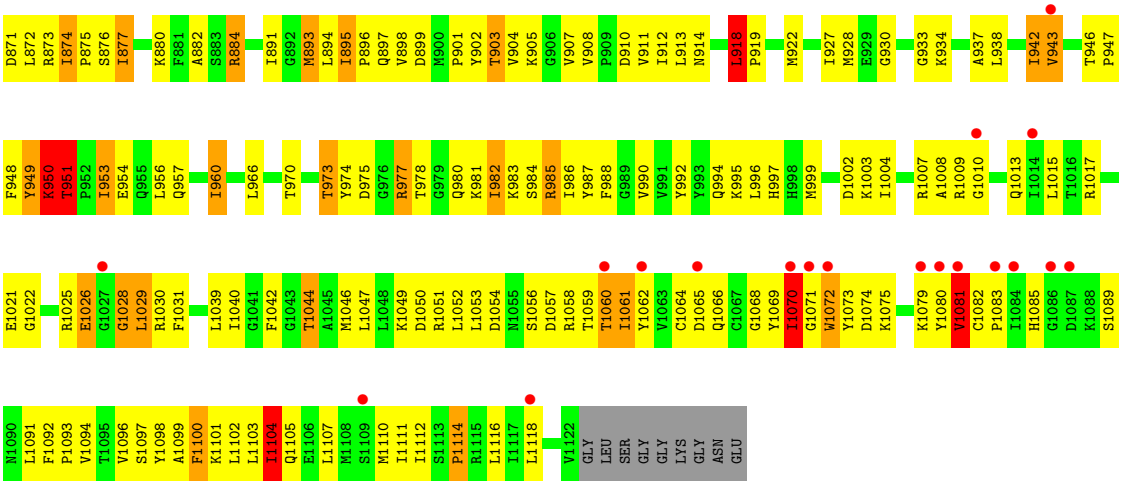


• Molecule 2: DNA-DIRECTED RNA POLYMERASE

Chain B:



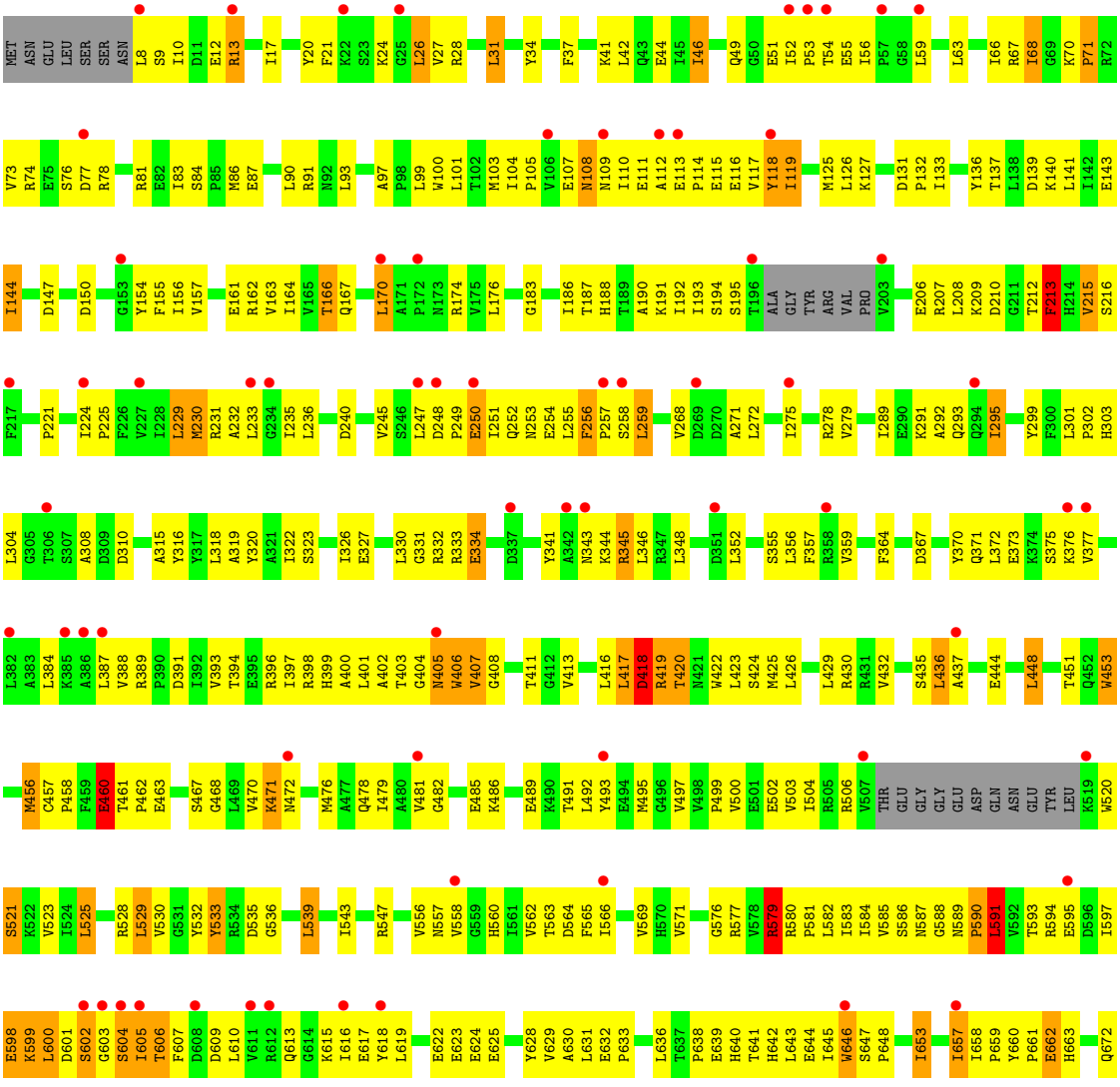


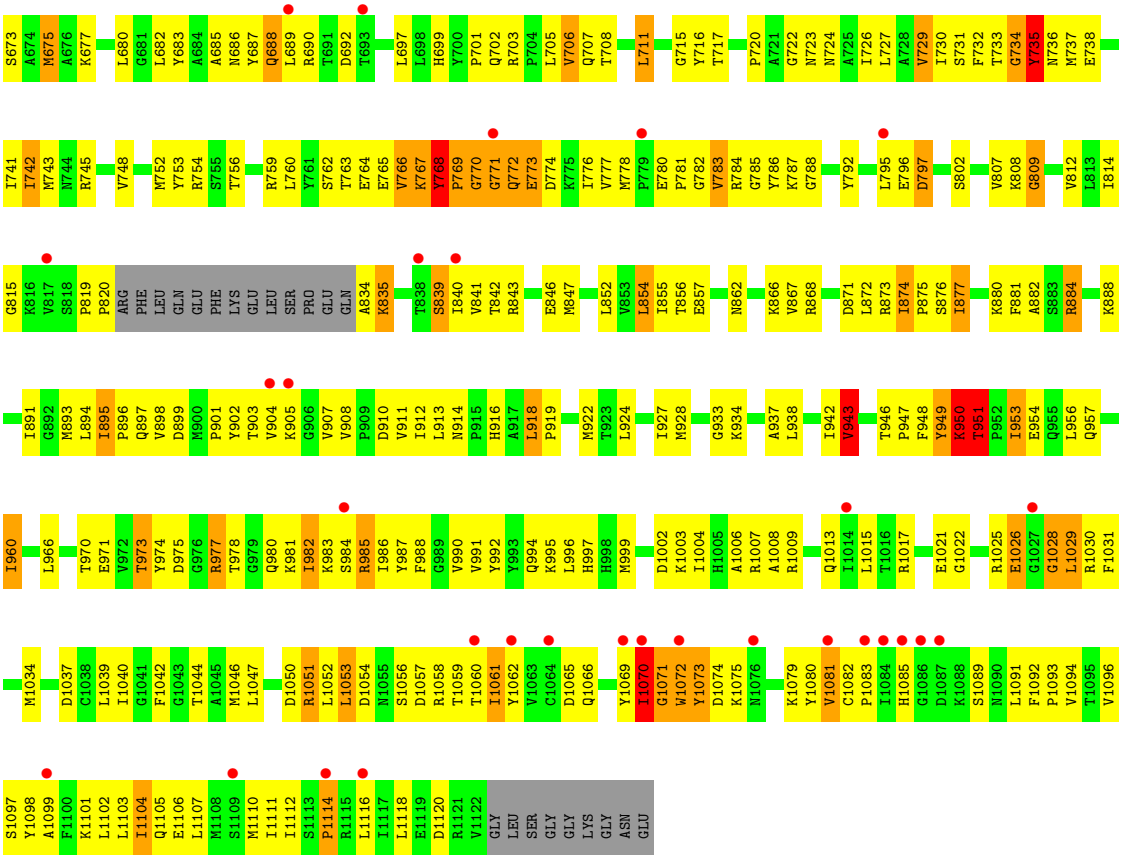


• Molecule 2: DNA-DIRECTED RNA POLYMERASE

Chain R:



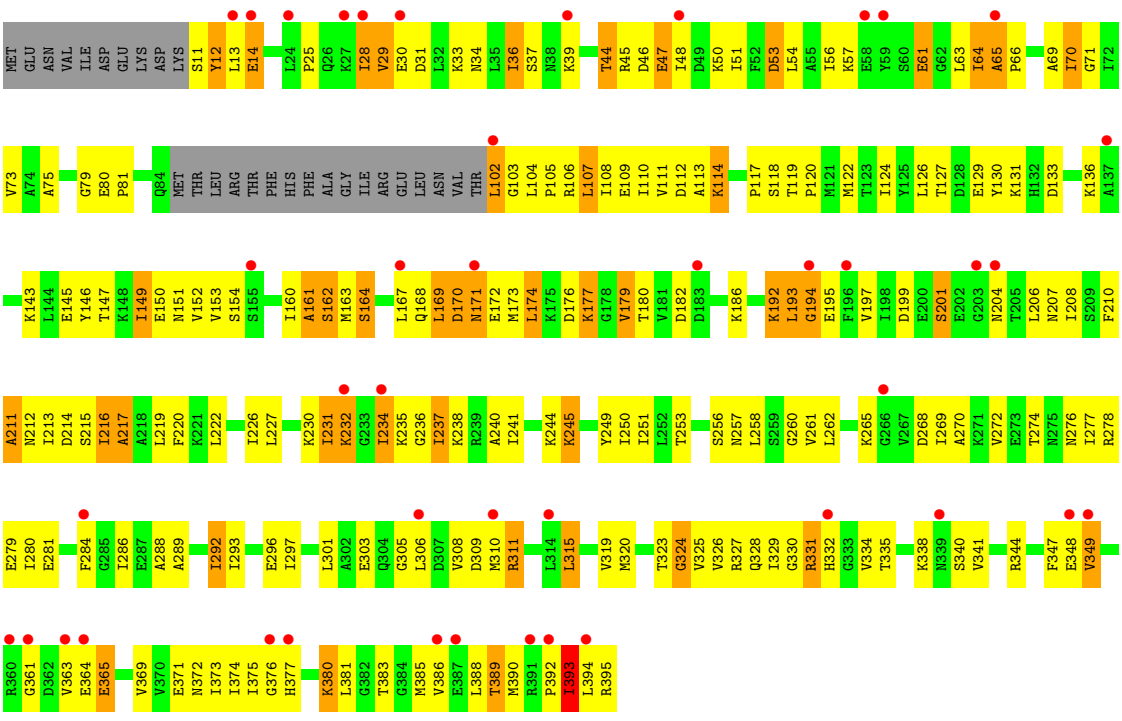




• Molecule 3: DNA-DIRECTED RNA POLYMERASE SUBUNIT A"

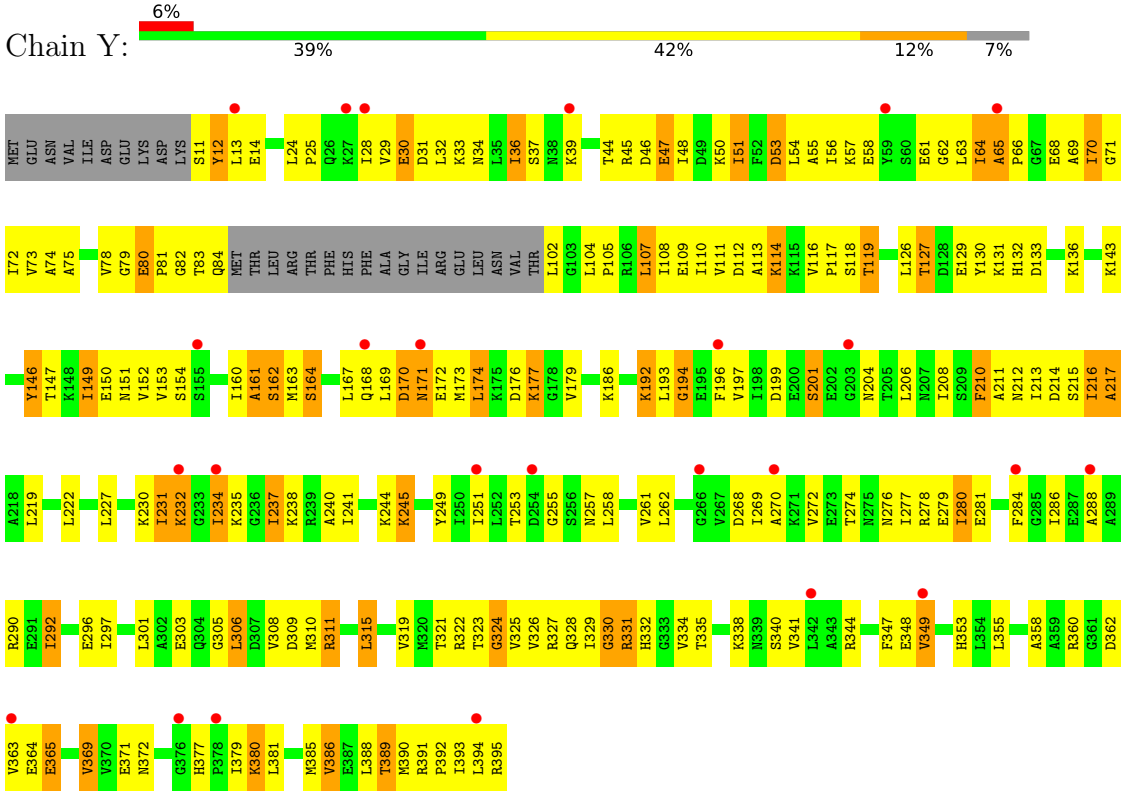
Chain C:





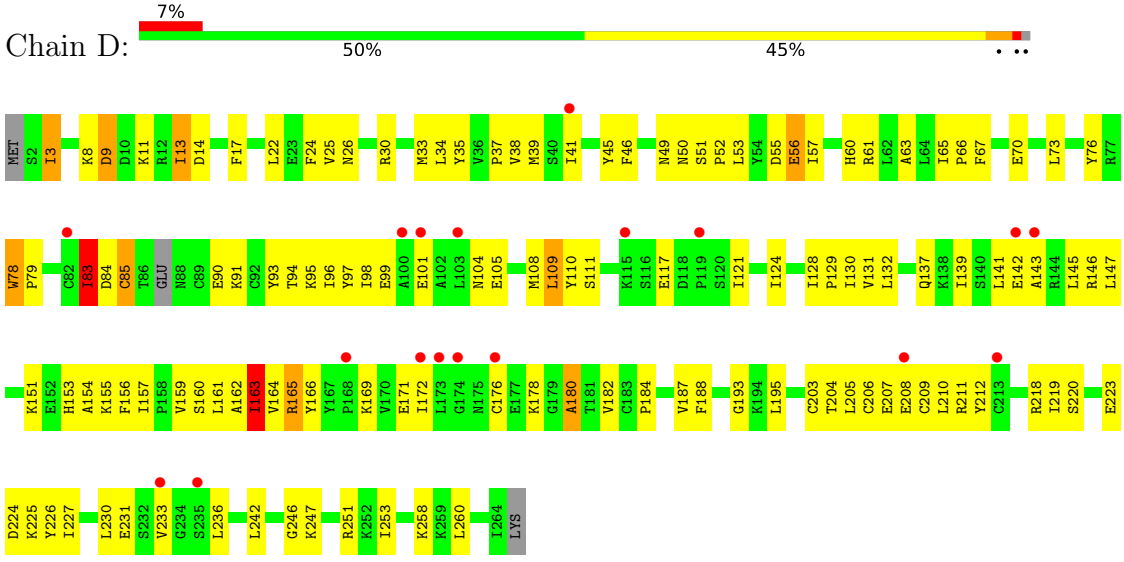
• Molecule 3: DNA-DIRECTED RNA POLYMERASE SUBUNIT A"

Chain Y:



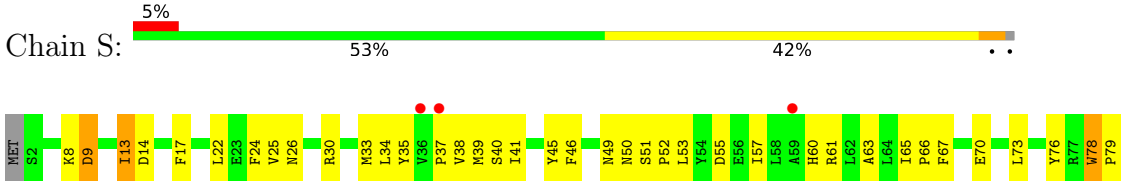
• Molecule 4: DNA-DIRECTED RNA POLYMERASE SUBUNIT D

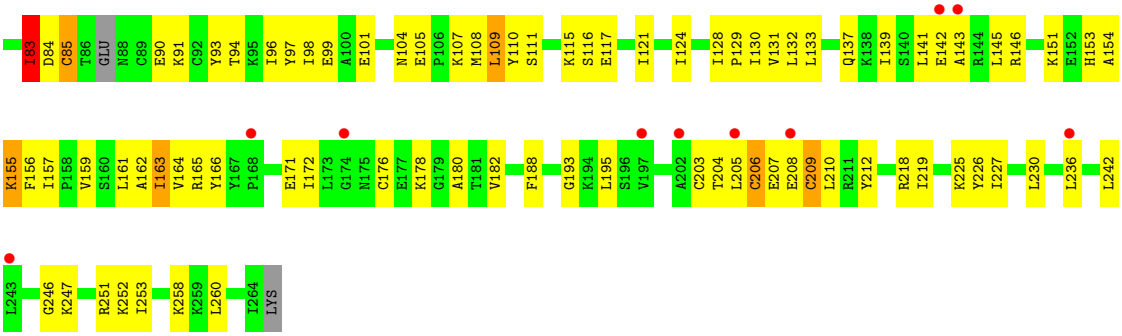
Chain D:



• Molecule 4: DNA-DIRECTED RNA POLYMERASE SUBUNIT D

Chain S:

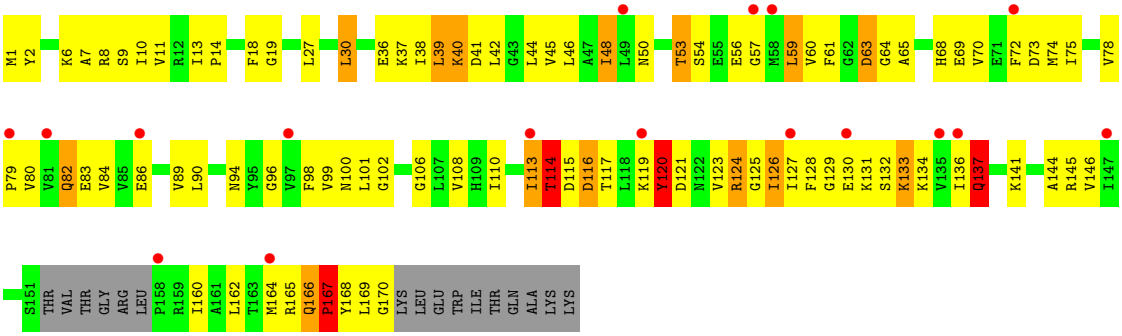




• Molecule 5: RNA POLYMERASE SUBUNIT 4

Chain E:

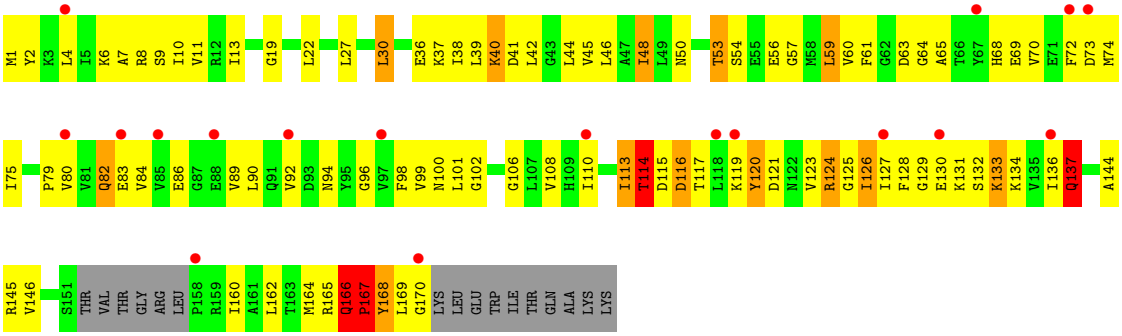




• Molecule 5: RNA POLYMERASE SUBUNIT 4

Chain T:

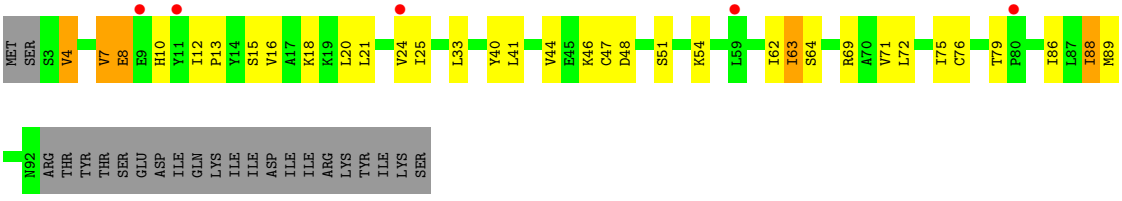




• Molecule 6: RNA POLYMERASE SUBUNIT 7

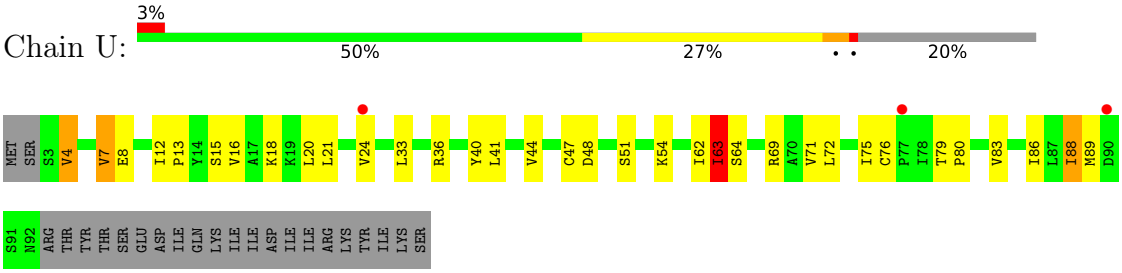
Chain F:





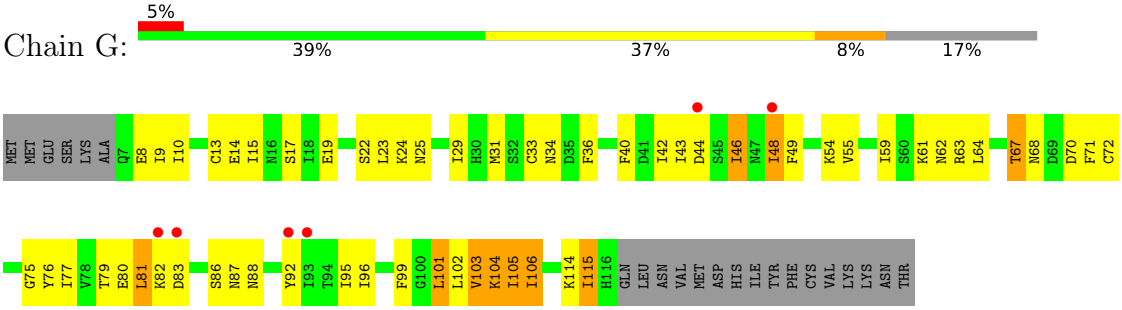
• Molecule 6: RNA POLYMERASE SUBUNIT 7

Chain U:



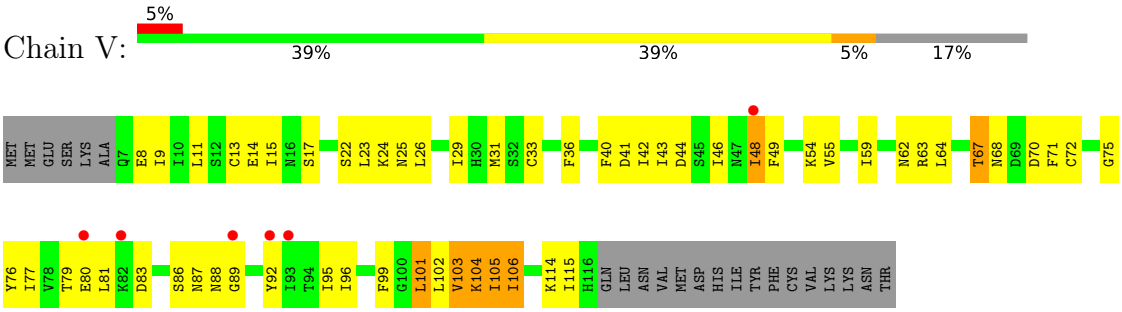
• Molecule 7: RNA POLYMERASE SUBUNIT 8

Chain G:



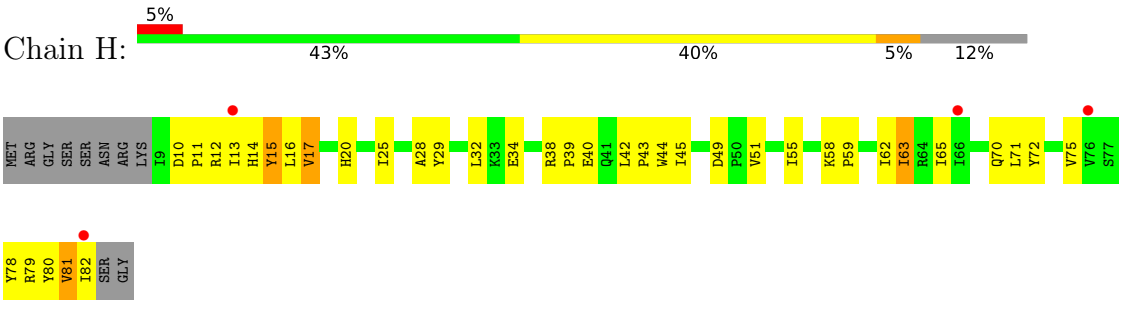
• Molecule 7: RNA POLYMERASE SUBUNIT 8

Chain V:



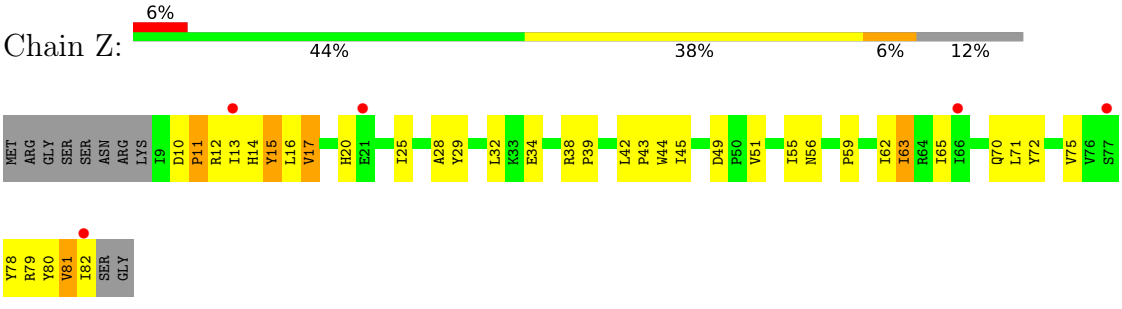
• Molecule 8: DNA-DIRECTED RNA POLYMERASE SUBUNIT H

Chain H:



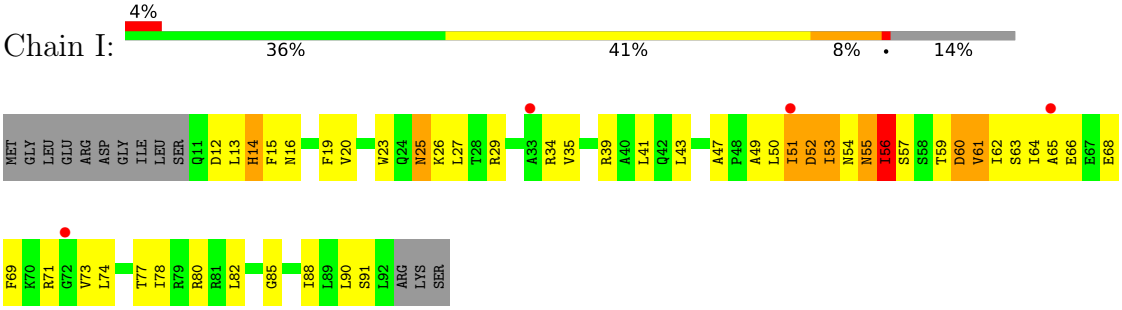
• Molecule 8: DNA-DIRECTED RNA POLYMERASE SUBUNIT H

Chain Z:



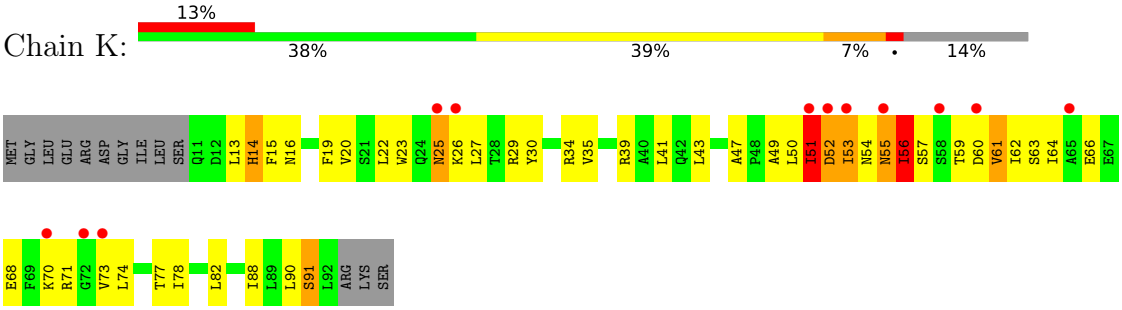
• Molecule 9: DNA-DIRECTED RNA POLYMERASE SUBUNIT K

Chain I:



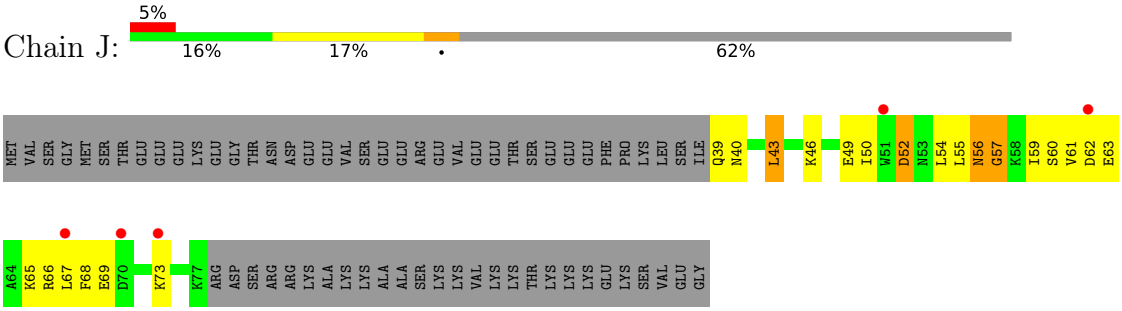
• Molecule 9: DNA-DIRECTED RNA POLYMERASE SUBUNIT K

Chain K:



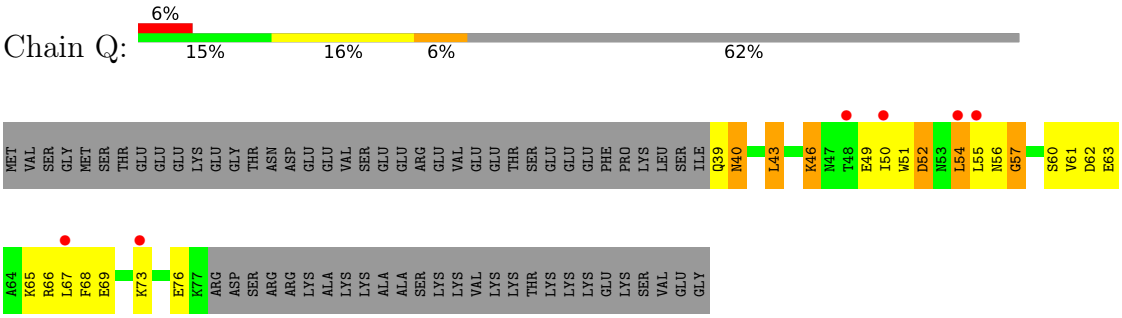
• Molecule 10: RNA POLYMERASE SUBUNIT 13

Chain J:



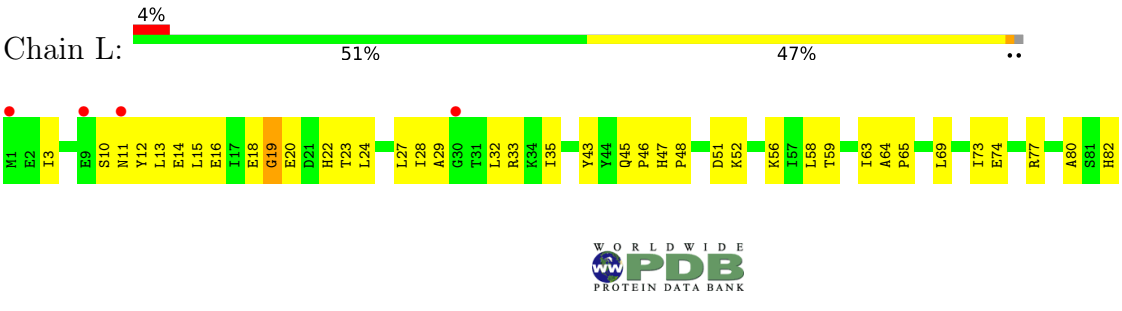
• Molecule 10: RNA POLYMERASE SUBUNIT 13

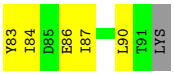
Chain Q:



• Molecule 11: DNA-DIRECTED RNA POLYMERASE SUBUNIT L

Chain L:

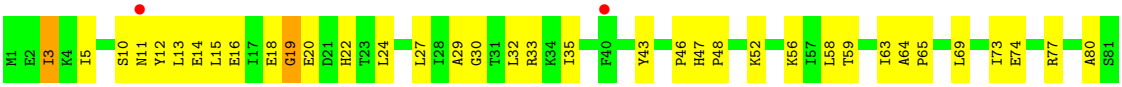


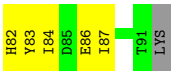


• Molecule 11: DNA-DIRECTED RNA POLYMERASE SUBUNIT L

Chain M:



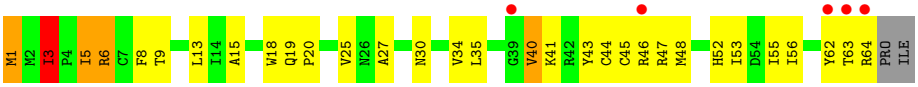




• Molecule 12: DNA-DIRECTED RNA POLYMERASE SUBUNIT N

Chain N:

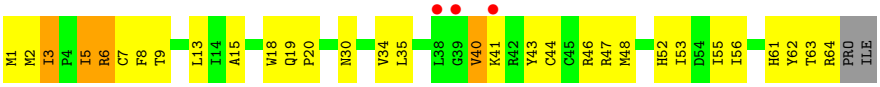




• Molecule 12: DNA-DIRECTED RNA POLYMERASE SUBUNIT N

Chain O:





• Molecule 13: RNA POLYMERASE SUBUNIT 12

Chain P:

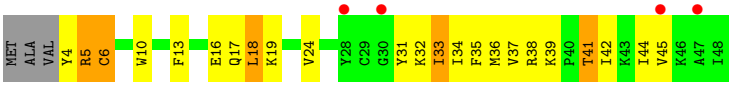




• Molecule 13: RNA POLYMERASE SUBUNIT 12

Chain X:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	129.33Å 195.91Å 212.45Å 90.00° 105.73° 90.00°	Depositor
Resolution (Å)	29.86 – 3.80 29.86 – 3.80	Depositor EDS
% Data completeness (in resolution range)	99.2 (29.86-3.80) 99.0 (29.86-3.80)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.94 (at 3.77Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.300 , 0.354 0.288 , 0.344	Depositor DCC
R_{free} test set	5001 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	79.6	Xtriage
Anisotropy	0.414	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 117.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.000 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	52472	wwPDB-VP
Average B, all atoms (Å^2)	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 21.01 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.6829e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1] Intensities estimated from amplitudes.

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

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, F3S, ZN

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	0.67	1/6762 (0.0%)	1.20	55/9151 (0.6%)
1	W	0.55	0/6762	0.95	23/9151 (0.3%)
2	B	0.68	1/8778 (0.0%)	1.17	55/11873 (0.5%)
2	R	0.56	0/8778	0.92	16/11873 (0.1%)
3	C	0.66	0/2869	1.18	18/3862 (0.5%)
3	Y	0.55	0/2869	0.96	10/3862 (0.3%)
4	D	0.51	0/2116	0.97	8/2859 (0.3%)
4	S	0.42	0/2116	0.77	3/2859 (0.1%)
5	E	0.55	1/1316 (0.1%)	1.07	3/1775 (0.2%)
5	T	0.51	1/1316 (0.1%)	0.87	3/1775 (0.2%)
6	F	0.38	0/709	0.94	1/961 (0.1%)
6	U	0.34	0/709	0.71	0/961
7	G	0.53	0/890	0.99	5/1194 (0.4%)
7	V	0.41	0/890	0.76	1/1194 (0.1%)
8	H	0.68	0/623	1.15	2/845 (0.2%)
8	Z	0.55	0/623	0.90	2/845 (0.2%)
9	I	2.28	5/662 (0.8%)	1.27	8/896 (0.9%)
9	K	2.19	5/662 (0.8%)	1.39	10/896 (1.1%)
10	J	0.54	1/336 (0.3%)	0.74	0/450
10	Q	0.68	2/336 (0.6%)	1.04	2/450 (0.4%)
11	L	0.55	0/717	1.10	4/968 (0.4%)
11	M	0.43	0/717	0.85	0/968
12	N	0.58	0/524	1.07	3/706 (0.4%)
12	O	0.48	0/524	0.85	0/706
13	P	0.59	0/378	1.04	1/507 (0.2%)
13	X	0.52	0/378	0.84	0/507
All	All	0.68	17/53360 (0.0%)	1.04	233/72094 (0.3%)

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
1	A	0	1
5	E	0	1
5	T	0	1
9	I	0	1
9	K	0	1
All	All	0	5

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
9	I	56	ILE	C-N	-43.24	0.82	1.33
9	K	55	ASN	C-N	37.07	1.84	1.33
9	K	56	ILE	C-N	-35.35	0.84	1.33
9	I	55	ASN	C-N	34.27	1.80	1.33
5	T	57	GLY	C-N	-8.94	1.21	1.33

The worst 5 of 233 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
9	I	55	ASN	O-C-N	12.11	138.56	122.82
9	I	56	ILE	O-C-N	-11.93	107.66	122.57
9	K	55	ASN	O-C-N	11.30	137.52	122.82
9	K	56	ILE	O-C-N	-10.66	109.24	122.57
9	I	56	ILE	CA-C-N	10.62	138.07	121.19

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	104	VAL	Mainchain
5	E	167	PRO	Peptide
9	I	56	ILE	Mainchain
9	K	56	ILE	Mainchain
5	T	167	PRO	Peptide

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	6620	0	6690	502	0
1	W	6620	0	6690	526	0
2	B	8615	0	8753	740	0
2	R	8615	0	8754	736	0
3	C	2845	0	3004	252	0
3	Y	2845	0	3004	255	0
4	D	2081	0	2122	134	0
4	S	2081	0	2123	124	0
5	E	1297	0	1348	128	0
5	T	1297	0	1348	128	0
6	F	701	0	708	49	0
6	U	701	0	708	47	0
7	G	878	0	886	57	0
7	V	878	0	886	60	0
8	H	609	0	640	45	0
8	Z	609	0	640	37	0
9	I	653	0	688	62	0
9	K	653	0	688	57	0
10	J	332	0	324	23	0
10	Q	332	0	324	30	0
11	L	707	0	739	37	0
11	M	707	0	739	34	0
12	N	514	0	528	33	0
12	O	514	0	528	37	0
13	P	369	0	396	49	0
13	X	369	0	396	52	0
14	A	2	0	0	0	0
14	B	3	0	0	0	0
14	N	1	0	0	0	0
14	O	1	0	0	0	0
14	P	1	0	0	0	0
14	R	3	0	0	0	0
14	W	2	0	0	3	0
14	X	1	0	0	0	0
15	A	1	0	0	0	0
15	W	1	0	0	0	0
16	D	7	0	0	4	0
16	S	7	0	0	10	0
All	All	52472	0	53654	3761	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 35.

The worst 5 of 3761 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)
9:I:55:ASN:C	9:I:56:ILE:N	1.80	1.37
9:K:55:ASN:C	9:K:56:ILE:N	1.84	1.34
9:I:56:ILE:O	9:I:57:SER:N	1.67	1.24
9:I:56:ILE:C	9:I:57:SER:CA	2.13	1.22
9:K:56:ILE:C	9:K:57:SER:CA	2.14	1.20

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	823/880 (94%)	669 (81%)	120 (15%)	34 (4%)	2	20
1	W	823/880 (94%)	672 (82%)	111 (14%)	40 (5%)	1	17
2	B	1077/1131 (95%)	887 (82%)	118 (11%)	72 (7%)	1	13
2	R	1077/1131 (95%)	881 (82%)	122 (11%)	74 (7%)	1	12
3	C	364/395 (92%)	278 (76%)	63 (17%)	23 (6%)	1	14
3	Y	364/395 (92%)	282 (78%)	54 (15%)	28 (8%)	1	11
4	D	258/265 (97%)	219 (85%)	31 (12%)	8 (3%)	3	25
4	S	258/265 (97%)	220 (85%)	30 (12%)	8 (3%)	3	25
5	E	160/180 (89%)	127 (79%)	21 (13%)	12 (8%)	1	11
5	T	160/180 (89%)	125 (78%)	24 (15%)	11 (7%)	1	12
6	F	88/113 (78%)	75 (85%)	11 (12%)	2 (2%)	5	30
6	U	88/113 (78%)	77 (88%)	9 (10%)	2 (2%)	5	30
7	G	108/132 (82%)	89 (82%)	15 (14%)	4 (4%)	2	21
7	V	108/132 (82%)	88 (82%)	16 (15%)	4 (4%)	2	21

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
8	H	72/84 (86%)	59 (82%)	9 (12%)	4 (6%)	1 15
8	Z	72/84 (86%)	59 (82%)	7 (10%)	6 (8%)	0 9
9	I	80/95 (84%)	64 (80%)	9 (11%)	7 (9%)	0 9
9	K	80/95 (84%)	64 (80%)	9 (11%)	7 (9%)	0 9
10	J	37/104 (36%)	27 (73%)	8 (22%)	2 (5%)	1 16
10	Q	37/104 (36%)	27 (73%)	8 (22%)	2 (5%)	1 16
11	L	89/92 (97%)	78 (88%)	9 (10%)	2 (2%)	5 30
11	M	89/92 (97%)	79 (89%)	8 (9%)	2 (2%)	5 30
12	N	62/66 (94%)	50 (81%)	10 (16%)	2 (3%)	3 24
12	O	62/66 (94%)	50 (81%)	9 (14%)	3 (5%)	2 17
13	P	43/48 (90%)	33 (77%)	7 (16%)	3 (7%)	1 12
13	X	43/48 (90%)	33 (77%)	7 (16%)	3 (7%)	1 12
All	All	6522/7170 (91%)	5312 (81%)	845 (13%)	365 (6%)	1 15

5 of 365 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	23	SER
1	A	103	ARG
1	A	108	GLU
1	A	683	GLU
1	A	737	VAL

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	723/766 (94%)	676 (94%)	47 (6%)	15 42
1	W	723/766 (94%)	672 (93%)	51 (7%)	13 39
2	B	937/975 (96%)	867 (92%)	70 (8%)	12 38
2	R	937/975 (96%)	866 (92%)	71 (8%)	12 37

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	C	316/341 (93%)	291 (92%)	25 (8%)	11 36
3	Y	316/341 (93%)	293 (93%)	23 (7%)	13 39
4	D	233/238 (98%)	226 (97%)	7 (3%)	36 57
4	S	233/238 (98%)	226 (97%)	7 (3%)	36 57
5	E	144/158 (91%)	134 (93%)	10 (7%)	14 40
5	T	144/158 (91%)	134 (93%)	10 (7%)	14 40
6	F	84/107 (78%)	80 (95%)	4 (5%)	23 47
6	U	84/107 (78%)	80 (95%)	4 (5%)	23 47
7	G	104/125 (83%)	95 (91%)	9 (9%)	9 33
7	V	104/125 (83%)	96 (92%)	8 (8%)	12 37
8	H	67/75 (89%)	65 (97%)	2 (3%)	36 57
8	Z	67/75 (89%)	65 (97%)	2 (3%)	36 57
9	I	72/83 (87%)	69 (96%)	3 (4%)	26 50
9	K	72/83 (87%)	68 (94%)	4 (6%)	19 45
10	J	37/96 (38%)	35 (95%)	2 (5%)	20 45
10	Q	37/96 (38%)	35 (95%)	2 (5%)	20 45
11	L	79/80 (99%)	78 (99%)	1 (1%)	61 71
11	M	79/80 (99%)	77 (98%)	2 (2%)	42 61
12	N	58/60 (97%)	52 (90%)	6 (10%)	7 27
12	O	58/60 (97%)	54 (93%)	4 (7%)	14 40
13	P	41/43 (95%)	38 (93%)	3 (7%)	13 39
13	X	41/43 (95%)	38 (93%)	3 (7%)	13 39
All	All	5790/6294 (92%)	5410 (93%)	380 (7%)	15 41

5 of 380 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	R	470	VAL
5	T	114	THR
2	R	591	LEU
2	R	943	VAL
7	V	106	ILE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 32 such sidechains are listed below:

Mol	Chain	Res	Type
1	W	698	ASN
1	W	847	GLN
3	C	189	ASN
2	B	1085	HIS
3	Y	189	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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 18 ligands modelled in this entry, 16 are monoatomic - leaving 2 for Mogul analysis.

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	$\#\|Z\| > 2$
16	F3S	S	1001	4	0,9,9	-	-	-		
16	F3S	D	1001	4	0,9,9	-	-	-		

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
16	F3S	S	1001	4	-	-	0/3/3/3
16	F3S	D	1001	4	-	-	0/3/3/3

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.

2 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	S	1001	F3S	10	0
16	D	1001	F3S	4	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
9	K	2
9	I	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	K	55:ASN	C	56:ILE	N	1.84
1	I	55:ASN	C	56:ILE	N	1.80
1	K	56:ILE	C	57:SER	N	0.84
1	I	56:ILE	C	57:SER	N	0.82

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	831/880 (94%)	0.69	77 (9%) 14 16	28, 92, 200, 308	0
1	W	831/880 (94%)	0.67	63 (7%) 20 18	46, 106, 212, 378	0
2	B	1085/1131 (95%)	0.74	91 (8%) 17 17	29, 88, 181, 293	0
2	R	1085/1131 (95%)	0.77	96 (8%) 15 16	55, 93, 190, 334	0
3	C	368/395 (93%)	0.86	43 (11%) 9 12	33, 108, 214, 329	0
3	Y	368/395 (93%)	0.71	25 (6%) 23 20	56, 115, 224, 338	0
4	D	262/265 (98%)	0.73	18 (6%) 23 20	59, 116, 219, 298	0
4	S	262/265 (98%)	0.58	13 (4%) 34 25	77, 134, 209, 265	0
5	E	164/180 (91%)	0.96	17 (10%) 11 14	70, 166, 254, 302	0
5	T	164/180 (91%)	0.96	18 (10%) 10 13	72, 164, 255, 300	0
6	F	90/113 (79%)	0.85	5 (5%) 30 23	98, 188, 257, 327	0
6	U	90/113 (79%)	0.81	3 (3%) 49 34	128, 184, 276, 323	0
7	G	110/132 (83%)	0.74	6 (5%) 30 23	66, 132, 226, 301	0
7	V	110/132 (83%)	0.68	6 (5%) 30 23	75, 145, 228, 295	0
8	H	74/84 (88%)	0.63	4 (5%) 31 24	53, 82, 153, 205	0
8	Z	74/84 (88%)	0.70	5 (6%) 23 20	83, 99, 171, 205	0
9	I	82/95 (86%)	0.64	4 (4%) 35 26	57, 91, 160, 195	0
9	K	82/95 (86%)	0.90	12 (14%) 6 9	35, 84, 181, 219	0
10	J	39/104 (37%)	1.06	5 (12%) 7 11	130, 183, 243, 268	0
10	Q	39/104 (37%)	1.23	6 (15%) 5 8	112, 162, 213, 265	0
11	L	91/92 (98%)	0.46	4 (4%) 39 28	65, 99, 194, 250	0
11	M	91/92 (98%)	0.29	2 (2%) 62 44	95, 121, 175, 247	0
12	N	64/66 (96%)	0.76	5 (7%) 19 17	59, 103, 177, 205	0
12	O	64/66 (96%)	0.42	3 (4%) 36 27	91, 112, 168, 192	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
13	P	45/48 (93%)	0.89	4 (8%) 15 16	71, 93, 189, 275	0
13	X	45/48 (93%)	0.87	4 (8%) 15 16	85, 105, 197, 249	0
All	All	6610/7170 (92%)	0.73	539 (8%) 17 17	28, 105, 217, 378	0

The worst 5 of 539 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	W	238	LYS	6.6
2	B	1072	TRP	6.1
1	W	24	VAL	5.9
2	R	1072	TRP	5.8
12	N	64	ARG	5.5

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

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

6.3 Carbohydrates (i)

There are no oligosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
15	MG	W	1882	1/1	0.41	0.31	246,246,246,246	0
15	MG	A	1882	1/1	0.42	0.31	246,246,246,246	0
14	ZN	B	2124	1/1	0.46	0.39	246,246,246,246	0
14	ZN	B	2125	1/1	0.52	0.25	246,246,246,246	0
14	ZN	R	2124	1/1	0.54	0.36	246,246,246,246	0
14	ZN	R	2125	1/1	0.67	0.27	246,246,246,246	0
14	ZN	W	1881	1/1	0.79	0.10	246,246,246,246	0
16	F3S	D	1001	7/7	0.82	0.34	232,235,241,245	0
14	ZN	R	2123	1/1	0.84	0.14	246,246,246,246	0
16	F3S	S	1001	7/7	0.85	0.26	247,251,257,260	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
14	ZN	B	2123	1/1	0.86	0.18	246,246,246,246	0
14	ZN	A	1881	1/1	0.89	0.07	246,246,246,246	0
14	ZN	A	1880	1/1	0.90	0.14	246,246,246,246	0
14	ZN	W	1880	1/1	0.94	0.09	246,246,246,246	0
14	ZN	P	1049	1/1	0.96	0.06	142,142,142,142	0
14	ZN	X	1049	1/1	0.96	0.06	157,157,157,157	0
14	ZN	O	1065	1/1	0.99	0.02	165,165,165,165	0
14	ZN	N	1065	1/1	1.00	0.02	130,130,130,130	0

6.5 Other polymers (i)

There are no such residues in this entry.