wwPDB X-ray Structure Validation Summary Report (i)

Jun 25, 2024 – 07:29 AM EDT

PDB ID	:	5X51
Title	:	RNA Polymerase II from Komagataella Pastoris (Type-3 crystal)
Authors	:	Ehara, H.; Umehara, T.; Sekine, S.; Yokoyama, S.
Deposited on	:	2017-02-14
Resolution	:	7.00 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:

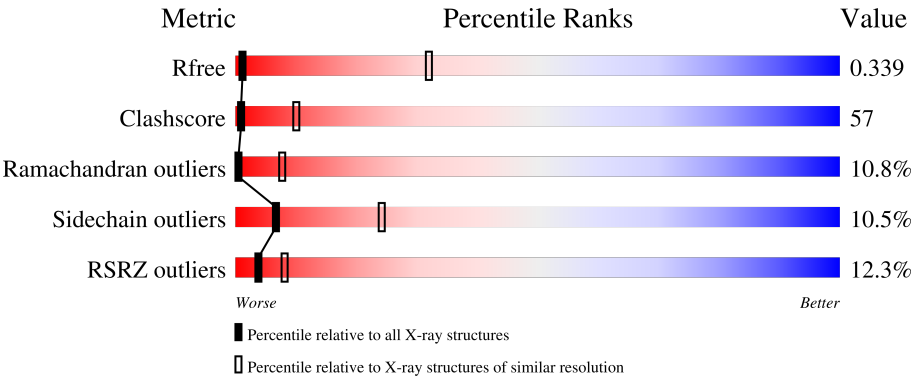
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 7.00 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}	130704	1004 (10.00-3.90)
Clashscore	141614	1069 (10.00-3.90)
Ramachandran outliers	138981	1002 (10.00-3.90)
Sidechain outliers	138945	1002 (10.00-3.86)
RSRZ outliers	127900	1004 (9.50-3.80)

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 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	1743	8% 27% 42% 9% • 21%
1	M	1743	10% 26% 42% 10% • 20%
2	B	1227	9% 26% 49% 11% • 13%
2	N	1227	11% 26% 49% 11% • 12%
3	C	304	4% 36% 39% 12% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	O	304	11% 36% 39% 11% 13%
4	D	186	18% 37% 40% 10% 12%
4	P	186	14% 35% 41% 10% 13%
5	E	214	21% 34% 55% 9%
5	Q	214	16% 34% 56% 9%
6	F	155	3% 15% 32% 6% 46%
6	R	155	5% 14% 32% 7% 46%
7	G	171	12% 36% 53% 11%
7	S	171	9% 39% 50% 11%
8	H	145	17% 30% 43% 15% 10%
8	T	145	20% 28% 45% 15% 10%
9	I	115	23% 36% 48% 12%
9	U	115	17% 35% 48% 13%
10	J	72	% 11% 58% 15% 14%
10	V	72	3% 14% 56% 15% 14%
11	K	118	11% 31% 52% 12%
11	W	118	14% 31% 52% 12%
12	L	73	18% 19% 26% 18% 37%
12	X	73	14% 23% 23% 16% 37%

2 Entry composition (i)

There are 13 unique types of molecules in this entry. The entry contains 56615 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 subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1384	Total	C	N	O	S	0	0	0
			10247	6499	1793	1901	54			
1	M	1386	Total	C	N	O	S	0	0	0
			10262	6507	1795	1906	54			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	1073	Total	C	N	O	S	0	0	0
			8182	5202	1425	1509	46			
2	N	1074	Total	C	N	O	S	0	0	0
			8190	5208	1426	1510	46			

- Molecule 3 is a protein called RNA polymerase II third largest subunit B44, part of central core.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	265	Total	C	N	O	S	0	0	0
			1863	1188	310	357	8			
3	O	265	Total	C	N	O	S	0	0	0
			1863	1188	310	357	8			

- Molecule 4 is a protein called RNA polymerase II subunit B32.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	163	Total	C	N	O	S	0	0	0
			1105	707	190	206	2			
4	P	162	Total	C	N	O	S	0	0	0
			1099	704	189	204	2			

- Molecule 5 is a protein called RNA polymerase subunit ABC27, common to RNA polymerases I, II, and III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	214	Total	C	N	O	S	0	0	0
			1638	1045	288	297	8			
5	Q	214	Total	C	N	O	S	0	0	0
			1638	1045	288	297	8			

- Molecule 6 is a protein called RNA polymerase subunit ABC23, common to RNA polymerases I, II, and III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	84	Total	C	N	O	S	0	0	0
			637	406	109	119	3			
6	R	84	Total	C	N	O	S	0	0	0
			637	406	109	119	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	171	Total	C	N	O	S	0	0	0
			1187	775	192	217	3			
7	S	171	Total	C	N	O	S	0	0	0
			1187	775	192	217	3			

- Molecule 8 is a protein called RNA polymerase subunit ABC14.5, common to RNA polymerases I, II, and III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	H	130	Total	C	N	O	S	0	0	0
			953	610	151	189	3			
8	T	130	Total	C	N	O	S	0	0	0
			953	610	151	189	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	I	112	Total	C	N	O	S	0	0	0
			847	531	149	156	11			
9	U	112	Total	C	N	O	S	0	0	0
			847	531	149	156	11			

- Molecule 10 is a protein called RNA polymerase subunit ABC10-beta, common to RNA polymerases I, II, and III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	J	62	Total	C	N	O	S	0	0	0
			487	320	85	76	6			
10	V	62	Total	C	N	O	S	0	0	0
			487	320	85	76	6			

- Molecule 11 is a protein called RNA polymerase II subunit B12.5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	K	113	Total	C	N	O	S	0	0	0
			826	539	138	147	2			
11	W	113	Total	C	N	O	S	0	0	0
			826	539	138	147	2			

- Molecule 12 is a protein called RNA polymerase subunit, found in RNA polymerase complexes I, II, and III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	L	46	Total	C	N	O	S	0	0	0
			319	196	64	55	4			
12	X	46	Total	C	N	O	S	0	0	0
			319	196	64	55	4			

- Molecule 13 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	A	2	Total	Zn	0	0
			2	2		
13	B	1	Total	Zn	0	0
			1	1		
13	C	1	Total	Zn	0	0
			1	1		
13	I	2	Total	Zn	0	0
			2	2		
13	J	1	Total	Zn	0	0
			1	1		
13	L	1	Total	Zn	0	0
			1	1		
13	M	2	Total	Zn	0	0
			2	2		
13	N	1	Total	Zn	0	0
			1	1		
13	O	1	Total	Zn	0	0
			1	1		

Continued on next page...

Continued from previous page...

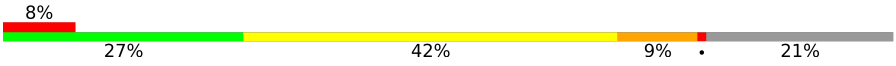
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
13	U	2	Total Zn 2 2	0	0
13	V	1	Total Zn 1 1	0	0
13	X	1	Total Zn 1 1	0	0

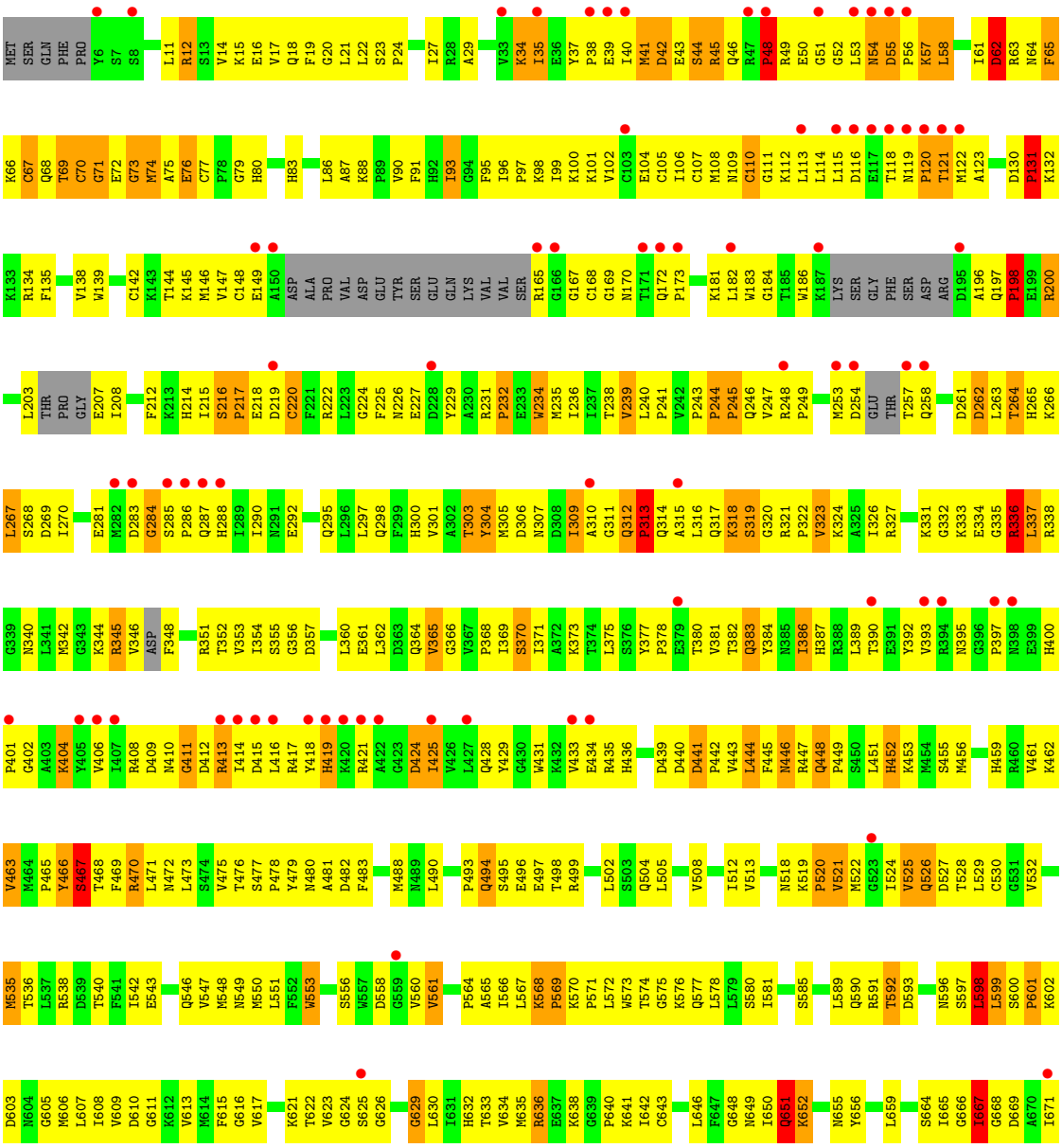
3 Residue-property plots (i)

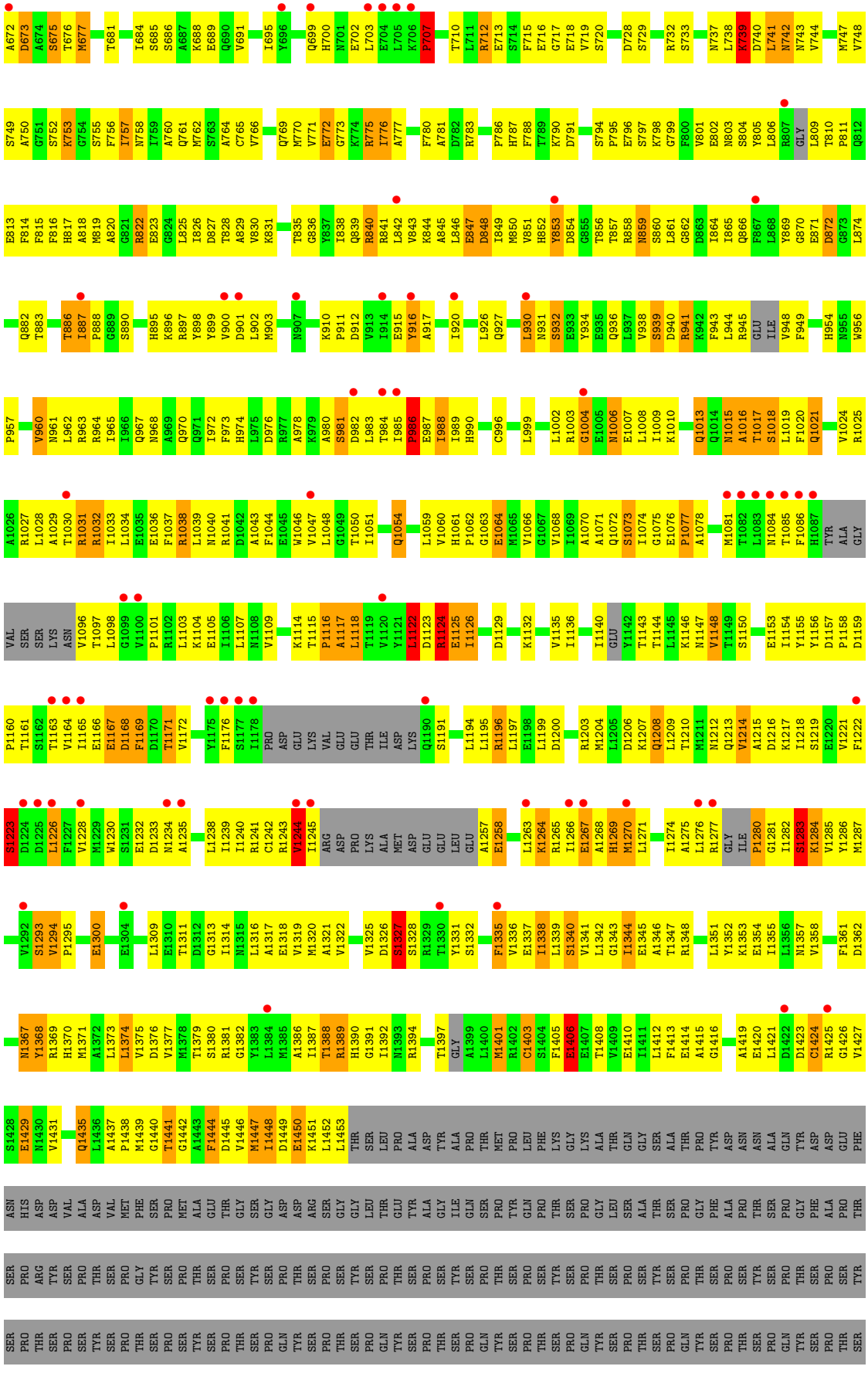
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density.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

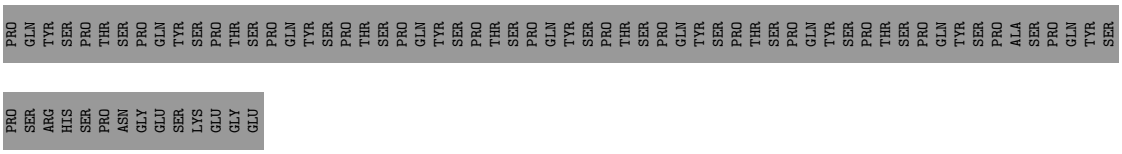
- Molecule 1: DNA-directed RNA polymerase subunit

Chain A:



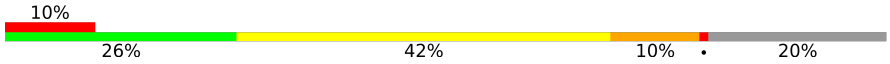


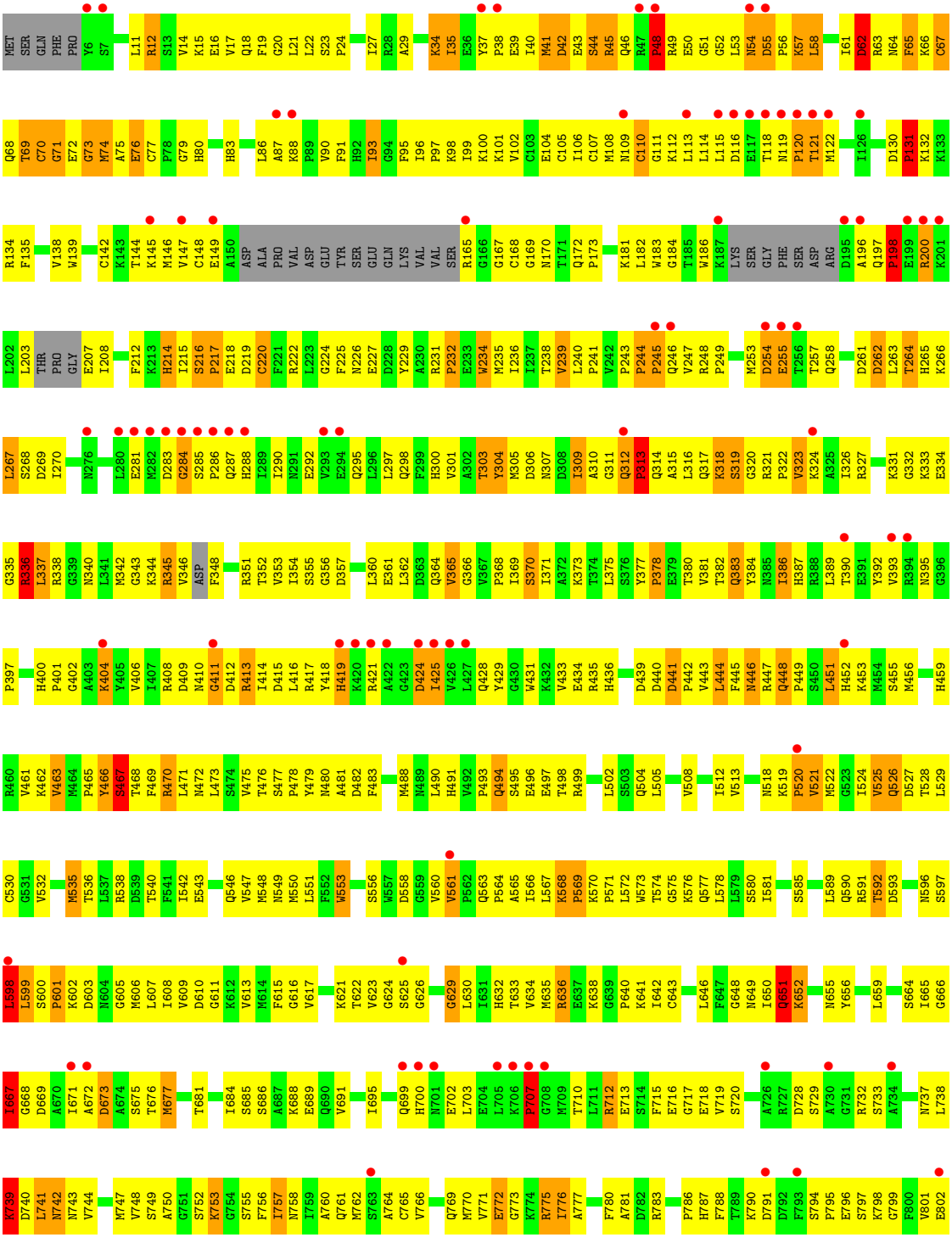


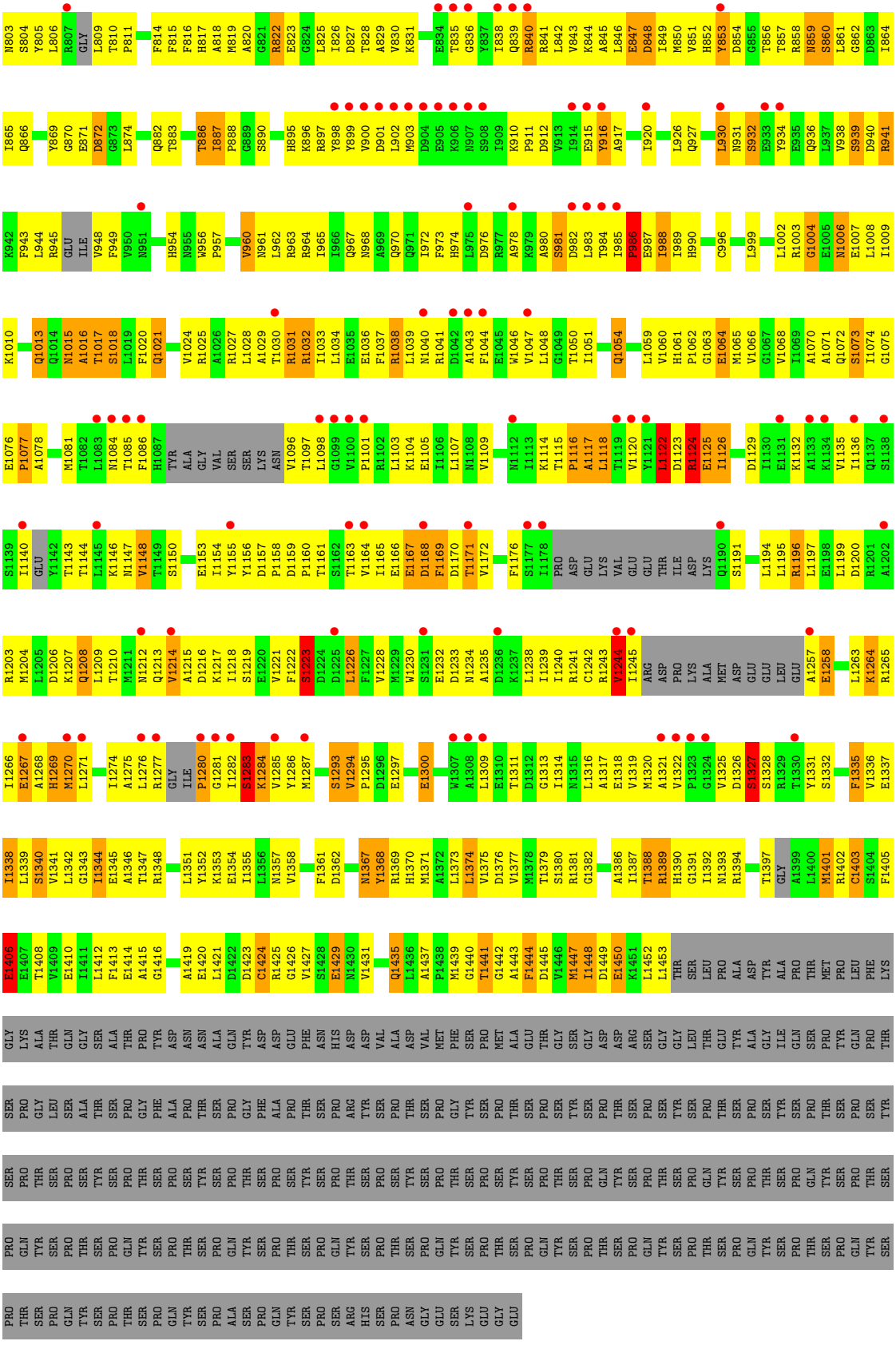


• Molecule 1: DNA-directed RNA polymerase subunit

Chain M:

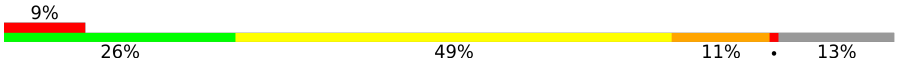


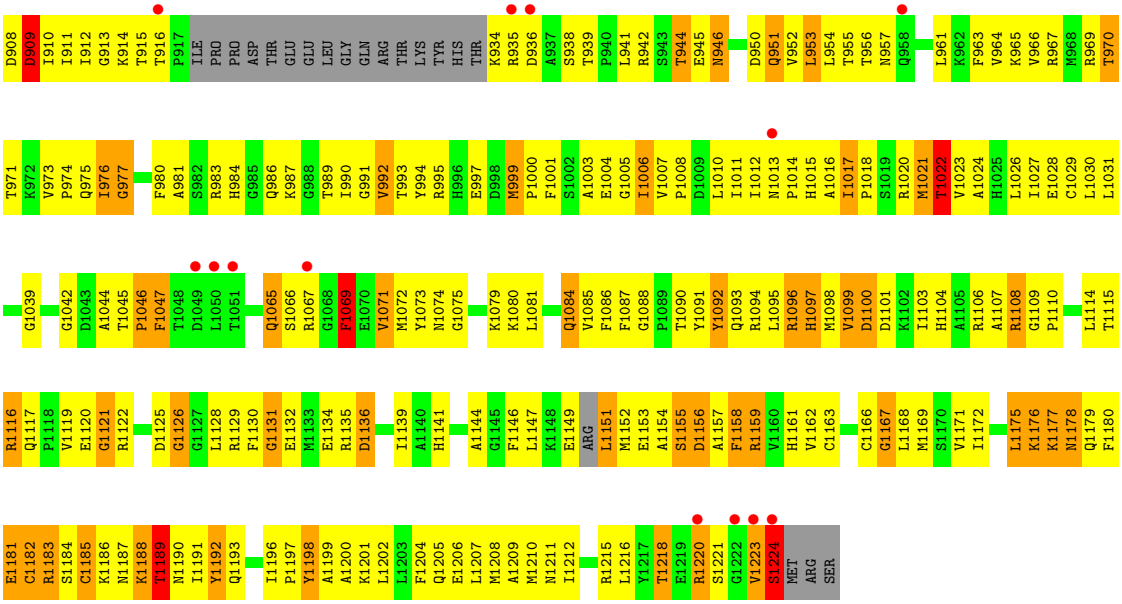




• Molecule 2: DNA-directed RNA polymerase subunit beta

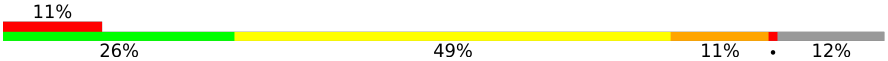
Chain B:

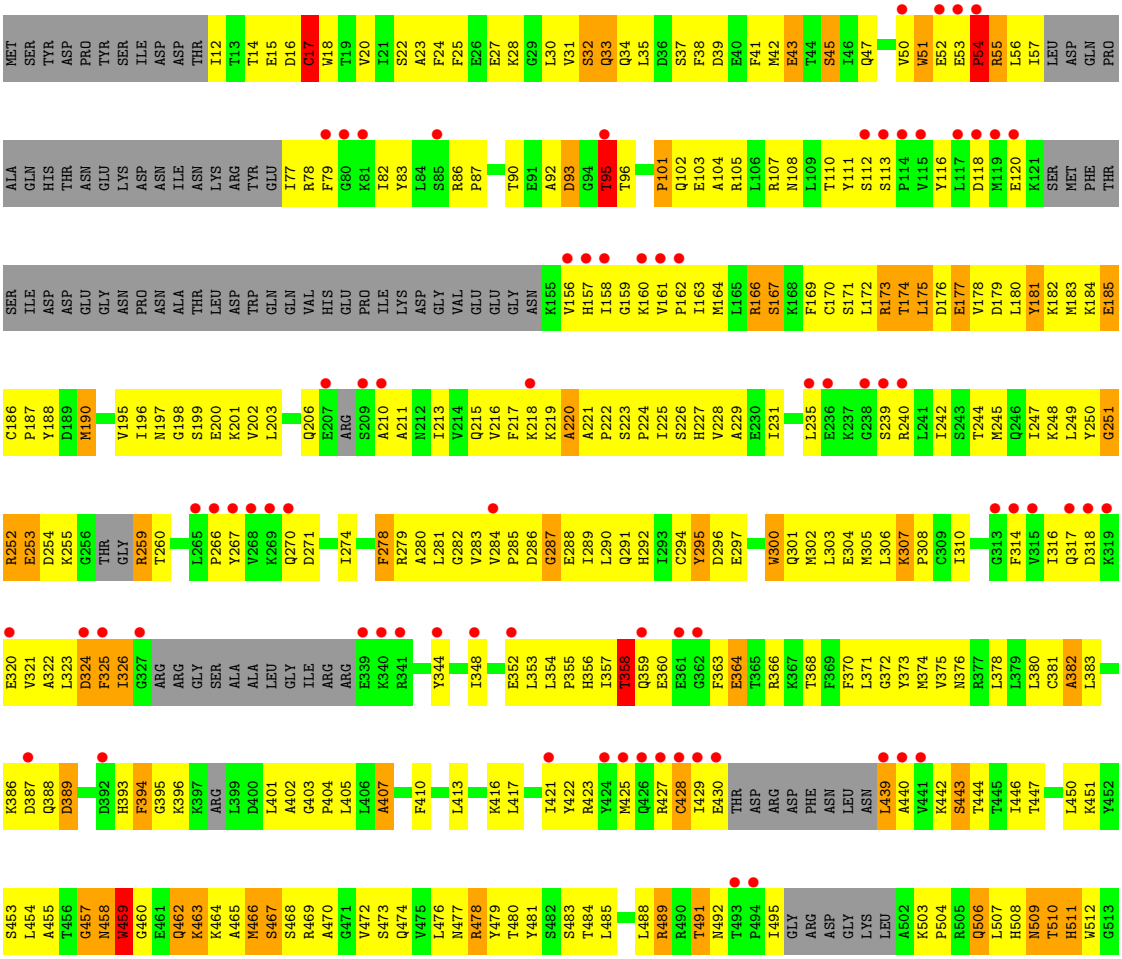


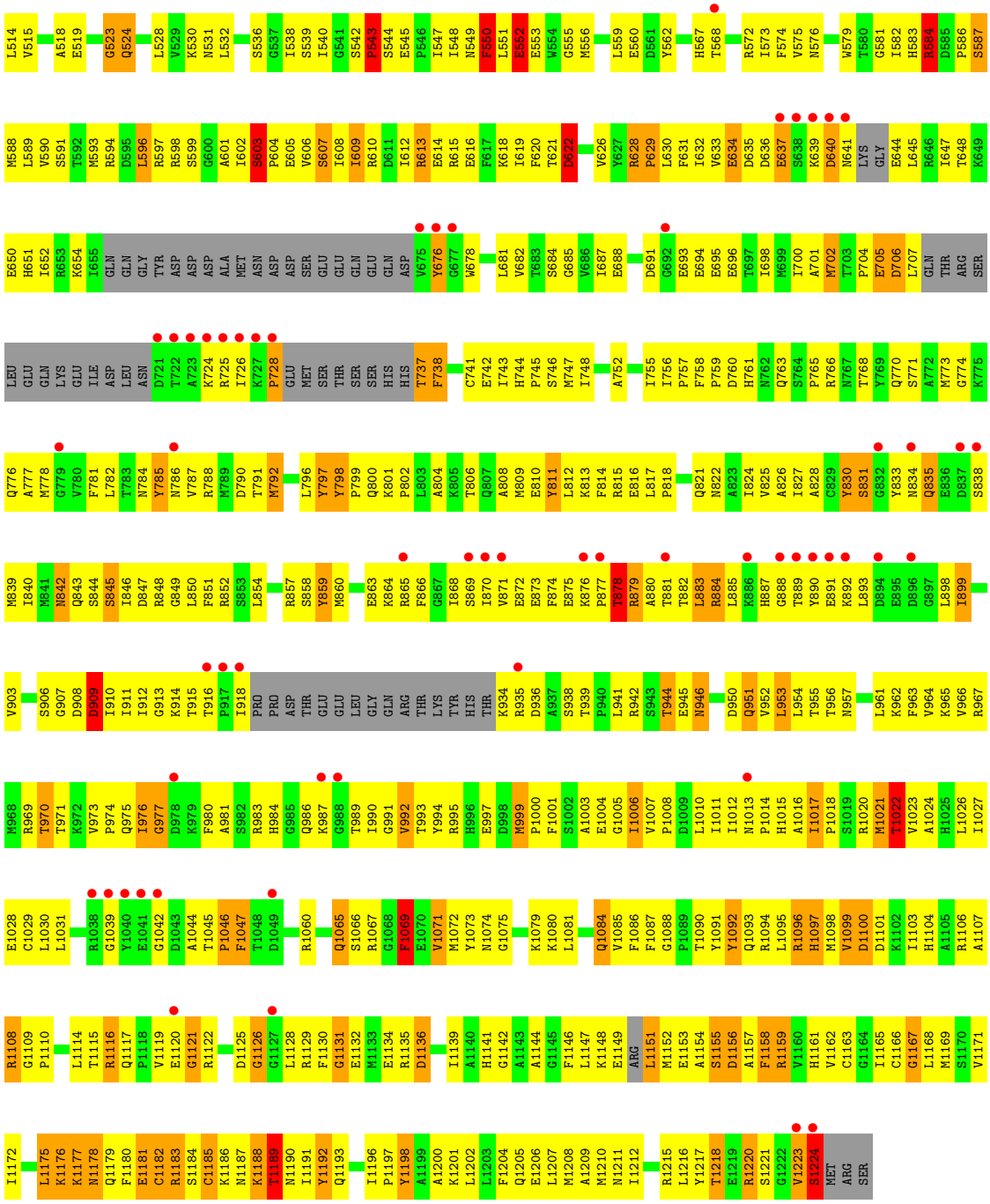


• Molecule 2: DNA-directed RNA polymerase subunit beta

Chain N:

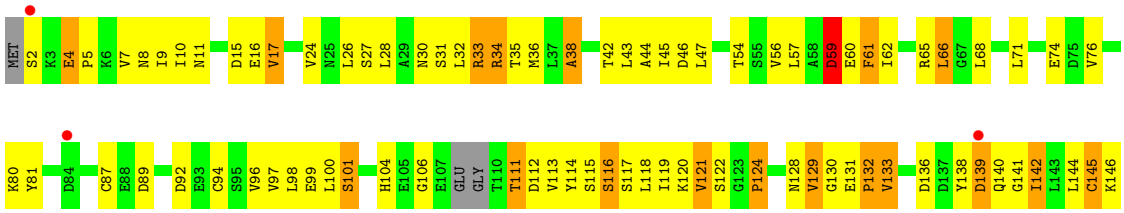


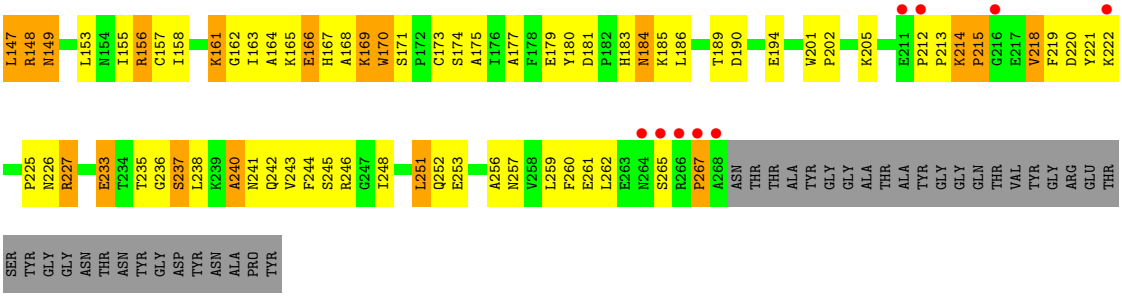




• Molecule 3: RNA polymerase II third largest subunit B44, part of central core

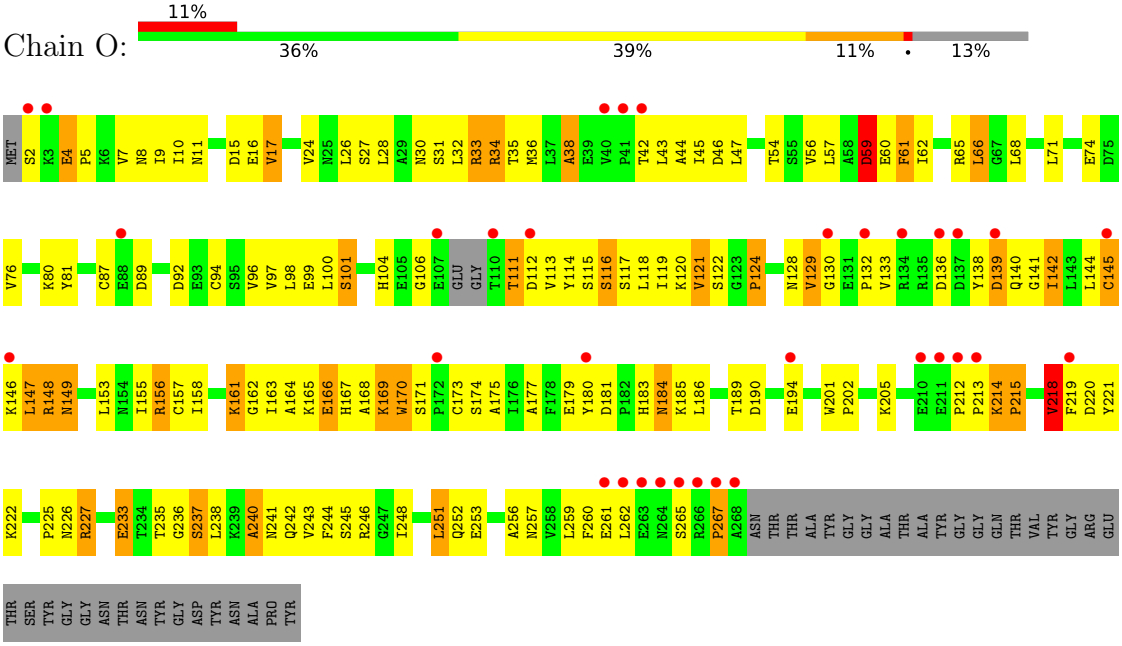
Chain C:





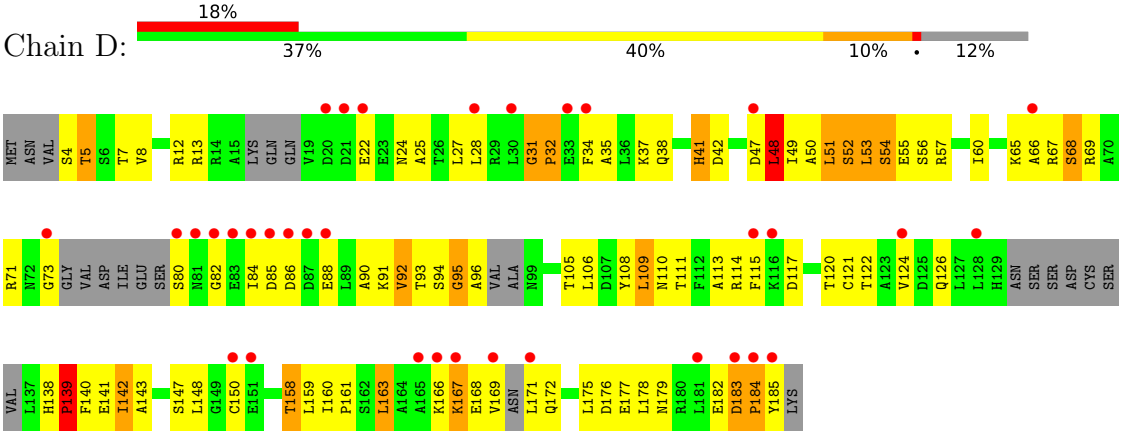
• Molecule 3: RNA polymerase II third largest subunit B44, part of central core

Chain O:



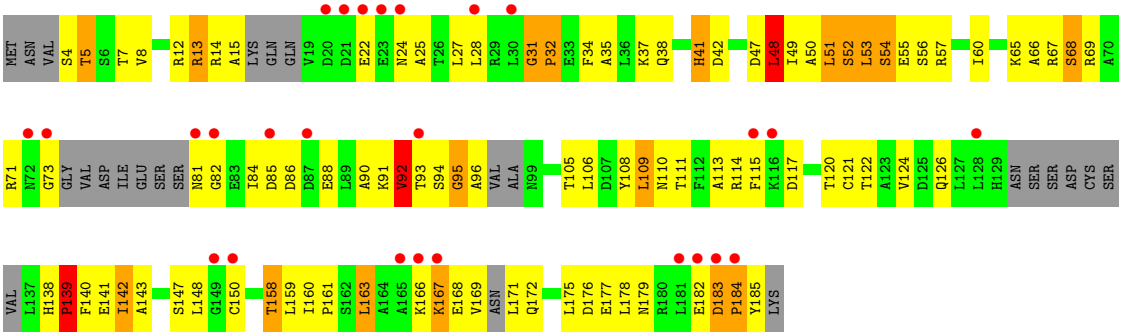
• Molecule 4: RNA polymerase II subunit B32

Chain D:



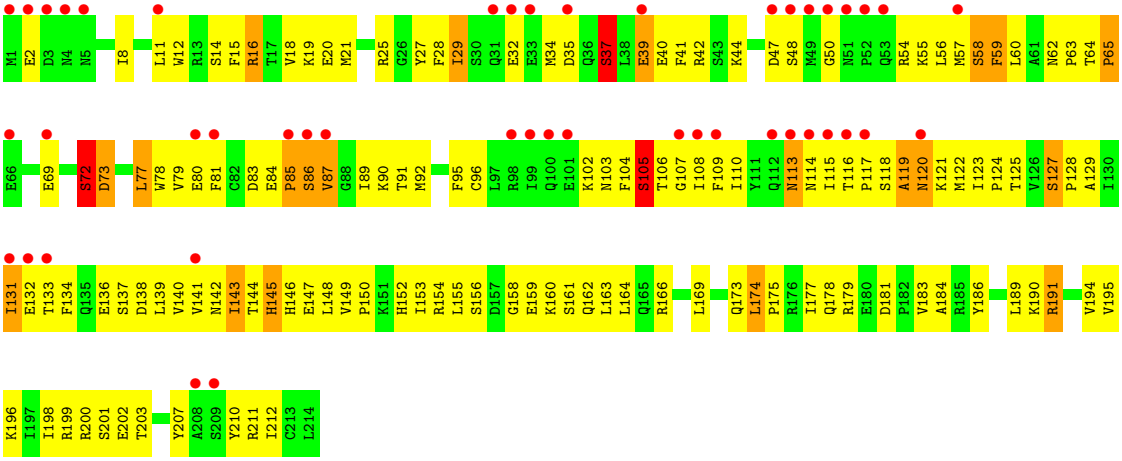
• Molecule 4: RNA polymerase II subunit B32

Chain P:



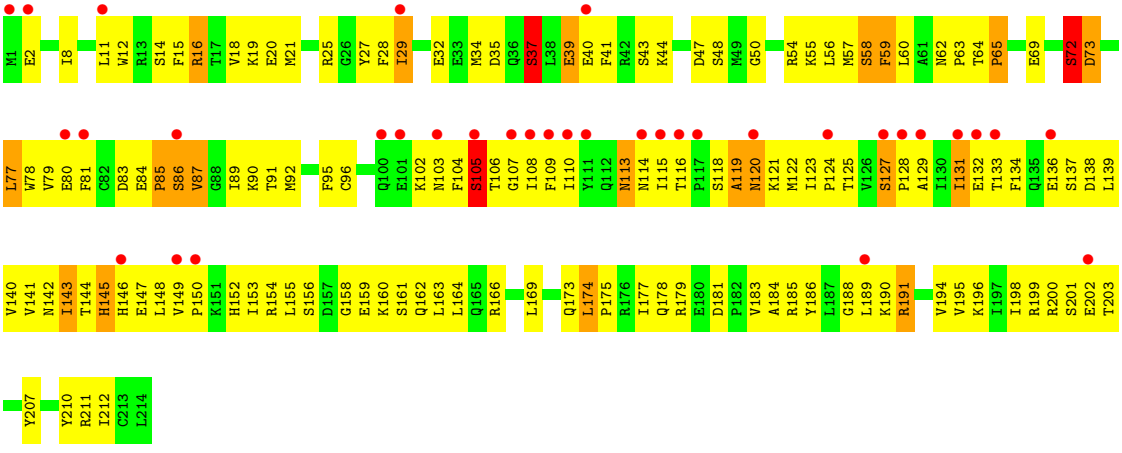
• Molecule 5: RNA polymerase subunit ABC27, common to RNA polymerases I, II, and III

Chain E:



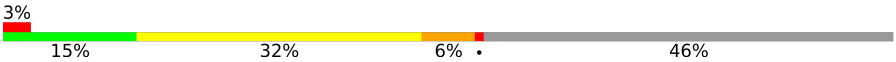
• Molecule 5: RNA polymerase subunit ABC27, common to RNA polymerases I, II, and III

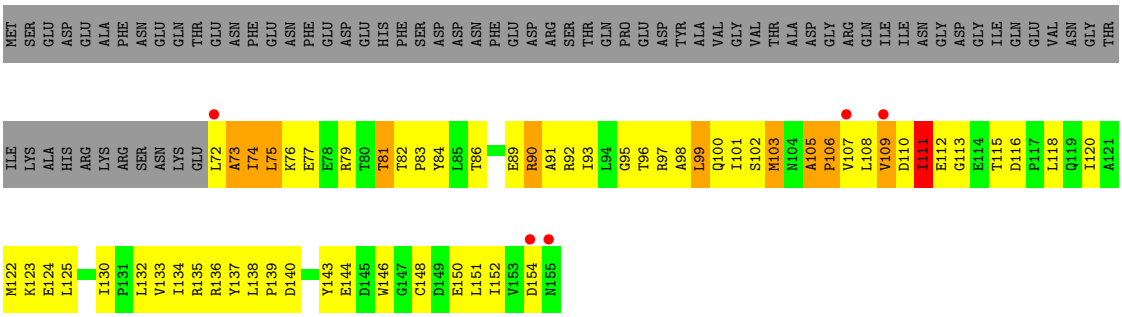
Chain Q:



• Molecule 6: RNA polymerase subunit ABC23, common to RNA polymerases I, II, and III

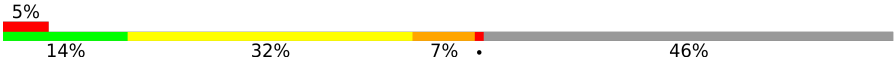
Chain F:

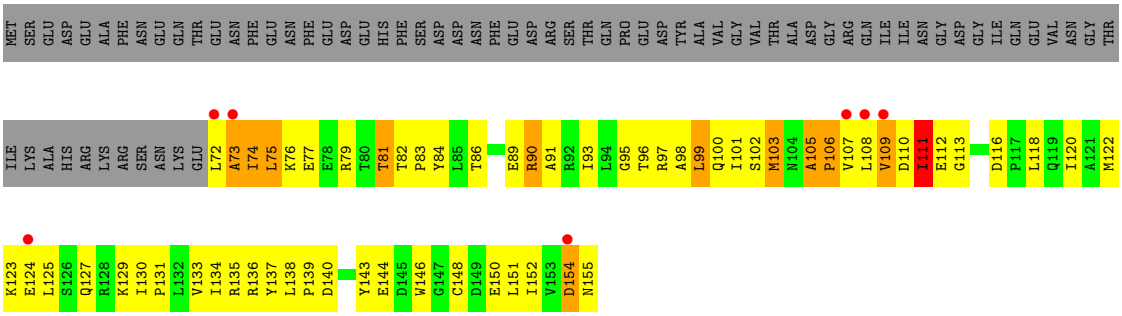




• Molecule 6: RNA polymerase subunit ABC23, common to RNA polymerases I, II, and III

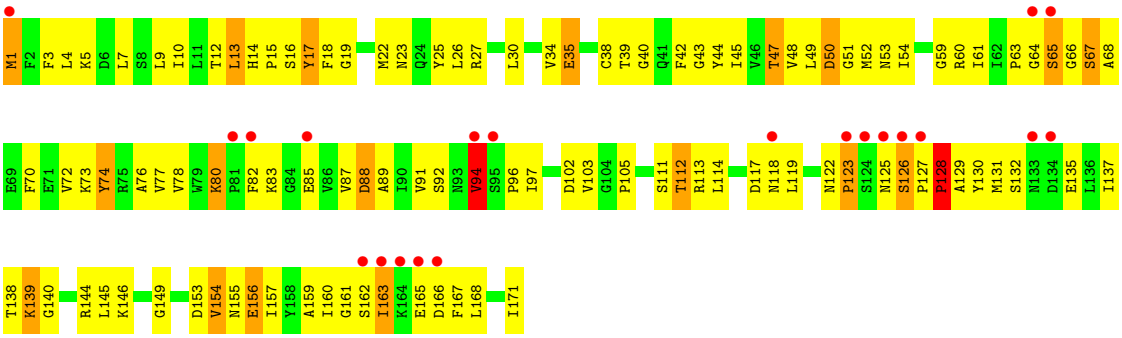
Chain R:





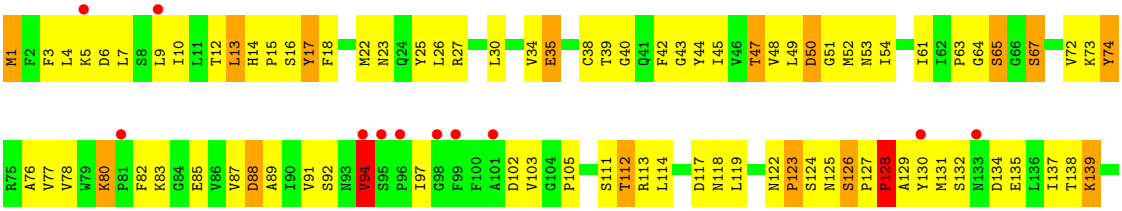
• Molecule 7: RNA polymerase II subunit

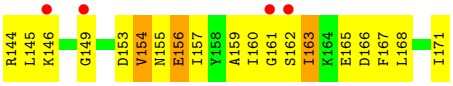
Chain G:



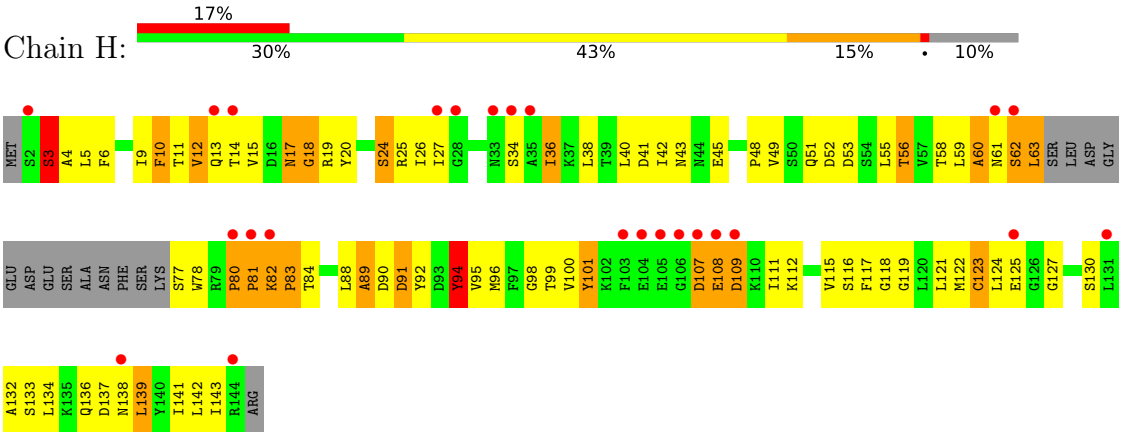
• Molecule 7: RNA polymerase II subunit

Chain S:

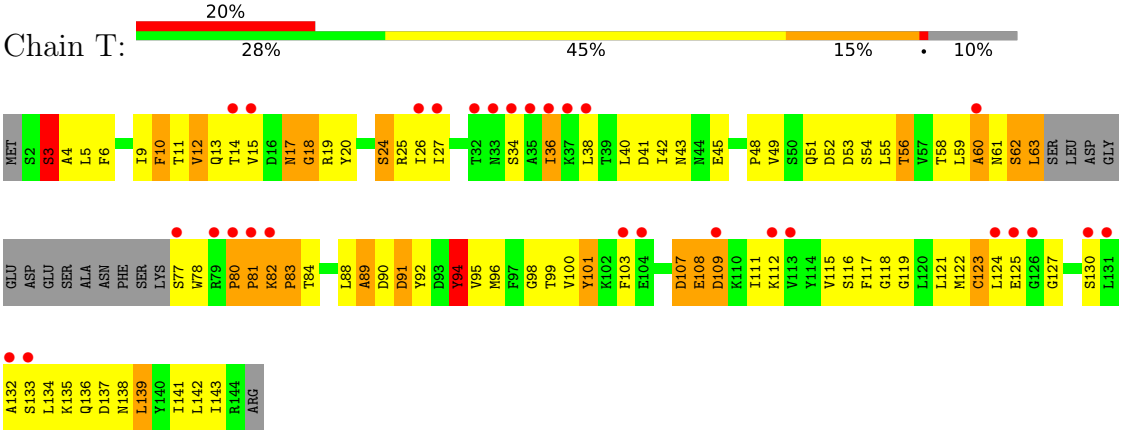




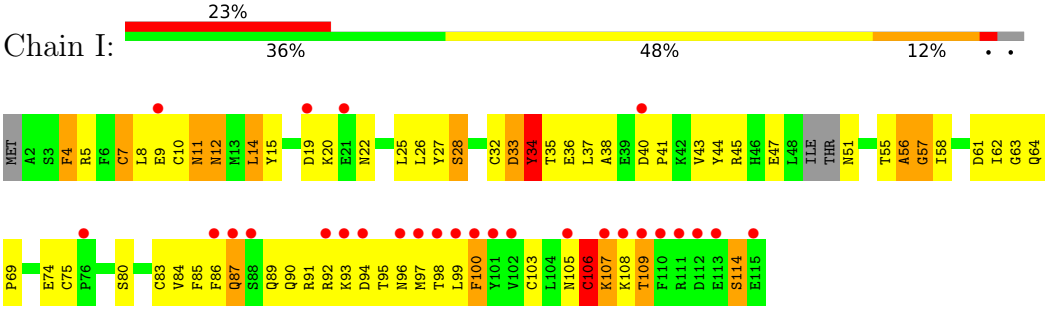
• Molecule 8: RNA polymerase subunit ABC14.5, common to RNA polymerases I, II, and III



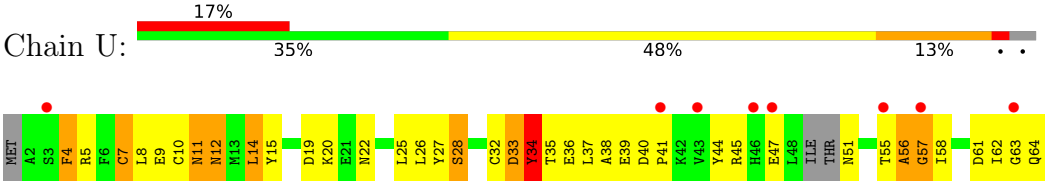
• Molecule 8: RNA polymerase subunit ABC14.5, common to RNA polymerases I, II, and III

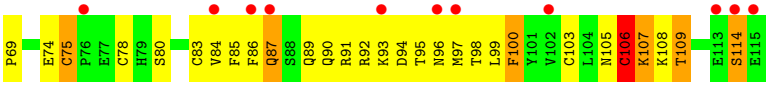


• Molecule 9: DNA-directed RNA polymerase subunit



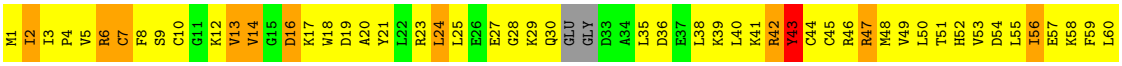
• Molecule 9: DNA-directed RNA polymerase subunit





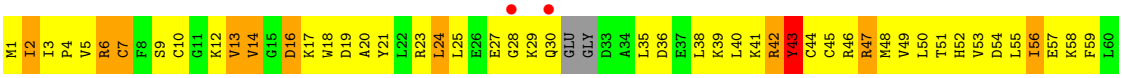
• Molecule 10: RNA polymerase subunit ABC10-beta, common to RNA polymerases I, II, and III

Chain J:



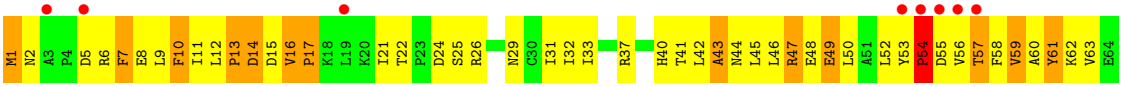
• Molecule 10: RNA polymerase subunit ABC10-beta, common to RNA polymerases I, II, and III

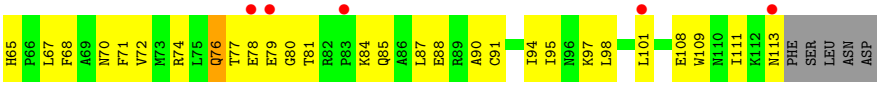
Chain V:



• Molecule 11: RNA polymerase II subunit B12.5

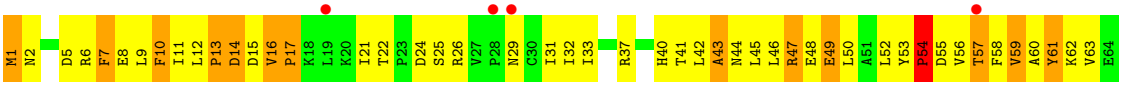
Chain K:

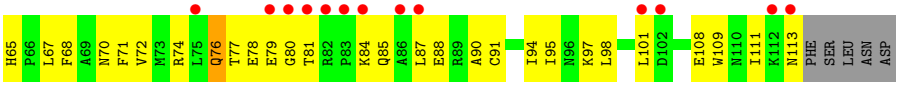




• Molecule 11: RNA polymerase II subunit B12.5

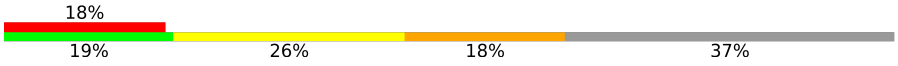
Chain W:

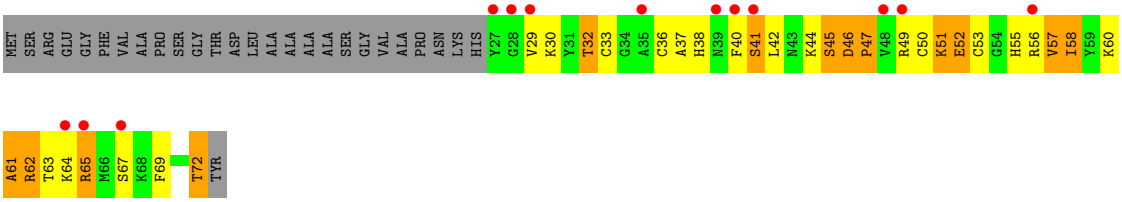




• Molecule 12: RNA polymerase subunit, found in RNA polymerase complexes I, II, and III

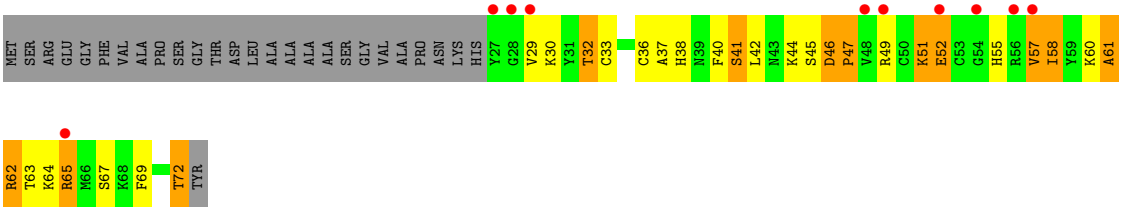
Chain L:





• Molecule 12: RNA polymerase subunit, found in RNA polymerase complexes I, II, and III

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	148.77Å 158.05Å 254.96Å 90.00° 92.54° 90.00°	Depositor
Resolution (Å)	33.27 – 7.00 49.58 – 7.00	Depositor EDS
% Data completeness (in resolution range)	99.5 (33.27-7.00) 99.4 (49.58-7.00)	Depositor EDS
R_{merge}	0.27	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.16 (at 6.68Å)	Xtriage
Refinement program	PHENIX dev_2614	Depositor
R, R_{free}	0.338 , 0.335 0.341 , 0.339	Depositor DCC
R_{free} test set	1886 reflections (10.06%)	wwPDB-VP
Wilson B-factor (Å^2)	221.0	Xtriage
Anisotropy	0.521	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 258.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.36$, $< L^2 > = 0.19$	Xtriage
Estimated twinning fraction	0.077 for h,-k,-l	Xtriage
F_o,F_c correlation	0.70	EDS
Total number of atoms	56615	wwPDB-VP
Average B, all atoms (Å^2)	261.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% 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, 0.333 respectively for untwinned datasets, and 0.375, 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: 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	0.68	27/10398 (0.3%)	0.81	17/14034 (0.1%)
1	M	0.68	28/10414 (0.3%)	0.81	18/14057 (0.1%)
2	B	0.69	14/8323 (0.2%)	0.77	4/11212 (0.0%)
2	N	0.69	15/8331 (0.2%)	0.77	4/11223 (0.0%)
3	C	0.70	6/1888 (0.3%)	0.83	8/2558 (0.3%)
3	O	0.71	6/1888 (0.3%)	0.83	8/2558 (0.3%)
4	D	0.87	6/1109 (0.5%)	0.76	4/1490 (0.3%)
4	P	0.83	5/1103 (0.5%)	0.76	4/1482 (0.3%)
5	E	0.73	7/1668 (0.4%)	0.69	3/2245 (0.1%)
5	Q	0.73	7/1668 (0.4%)	0.69	4/2245 (0.2%)
6	F	0.84	1/646 (0.2%)	0.82	2/873 (0.2%)
6	R	0.84	1/646 (0.2%)	0.82	2/873 (0.2%)
7	G	0.94	7/1207 (0.6%)	0.80	1/1629 (0.1%)
7	S	0.94	7/1207 (0.6%)	0.80	1/1629 (0.1%)
8	H	1.39	8/968 (0.8%)	0.75	2/1310 (0.2%)
8	T	1.39	8/968 (0.8%)	0.75	2/1310 (0.2%)
9	I	0.92	4/861 (0.5%)	0.70	0/1159
9	U	0.92	4/861 (0.5%)	0.70	0/1159
10	J	0.55	0/495	0.80	0/664
10	V	0.55	0/495	0.80	0/664
11	K	0.65	2/842 (0.2%)	0.85	5/1139 (0.4%)
11	W	0.64	2/842 (0.2%)	0.85	5/1139 (0.4%)
12	L	0.83	3/321 (0.9%)	0.96	2/425 (0.5%)
12	X	0.83	3/321 (0.9%)	0.96	2/425 (0.5%)
All	All	0.75	171/57470 (0.3%)	0.79	98/77502 (0.1%)

The worst 5 of 171 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	H	24	SER	CB-OG	30.84	1.82	1.42
8	T	24	SER	CB-OG	30.64	1.82	1.42

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	N	17	CYS	CB-SG	-23.56	1.42	1.82
2	B	17	CYS	CB-SG	-23.49	1.42	1.82
2	N	428	CYS	CB-SG	-22.71	1.43	1.82

The worst 5 of 98 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
12	X	29	VAL	CA-CB-CG1	9.64	125.36	110.90
12	L	29	VAL	CA-CB-CG2	9.62	125.33	110.90
2	N	682	VAL	CA-CB-CG1	7.27	121.81	110.90
2	B	682	VAL	CA-CB-CG2	7.27	121.81	110.90
11	W	54	PRO	N-CA-CB	7.01	111.71	103.30

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	10247	0	9733	1266	2
1	M	10262	0	9743	1348	1
2	B	8182	0	7871	990	0
2	N	8190	0	7882	1121	0
3	C	1863	0	1645	198	6
3	O	1863	0	1645	190	5
4	D	1105	0	964	99	5
4	P	1099	0	958	117	3
5	E	1638	0	1551	159	0
5	Q	1638	0	1551	159	0
6	F	637	0	620	97	0
6	R	637	0	620	94	0
7	G	1187	0	1092	224	5
7	S	1187	0	1092	149	13
8	H	953	0	851	118	4
8	T	953	0	851	124	2
9	I	847	0	756	79	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	U	847	0	756	83	7
10	J	487	0	492	120	0
10	V	487	0	492	114	0
11	K	826	0	723	89	5
11	W	826	0	723	88	6
12	L	319	0	287	31	8
12	X	319	0	287	27	6
13	A	2	0	0	0	0
13	B	1	0	0	0	0
13	C	1	0	0	0	0
13	I	2	0	0	0	0
13	J	1	0	0	0	0
13	L	1	0	0	0	0
13	M	2	0	0	0	0
13	N	1	0	0	1	0
13	O	1	0	0	0	0
13	U	2	0	0	0	0
13	V	1	0	0	0	0
13	X	1	0	0	0	0
All	All	56615	0	53185	6279	39

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 57.

The worst 5 of 6279 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:M:1397:THR:HG21	1:M:1401:MET:SD	1.33	1.67
2:N:51:TRP:CB	2:N:51:TRP:CG	1.81	1.64
1:M:521:VAL:CG2	1:M:521:VAL:CB	1.76	1.62
2:B:51:TRP:CG	2:B:51:TRP:CB	1.81	1.60
1:M:239:VAL:CG1	1:M:239:VAL:CB	1.80	1.60

The worst 5 of 39 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)
4:D:73:GLY:C	12:L:49:ARG:NH2[2_655]	1.24	0.96
4:D:73:GLY:O	12:L:49:ARG:NH2[2_655]	1.30	0.90
7:S:126:SER:O	9:U:38:ALA:O[1_455]	1.32	0.88

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:139:ASP:CB	11:W:113:ASN:C[2_646]	1.49	0.71
11:K:113:ASN:C	3:O:139:ASP:CB[2_646]	1.53	0.67

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	1356/1743 (78%)	945 (70%)	279 (21%)	132 (10%)	0 10
1	M	1360/1743 (78%)	947 (70%)	279 (20%)	134 (10%)	0 9
2	B	1043/1227 (85%)	711 (68%)	210 (20%)	122 (12%)	0 6
2	N	1044/1227 (85%)	712 (68%)	210 (20%)	122 (12%)	0 6
3	C	261/304 (86%)	182 (70%)	56 (22%)	23 (9%)	1 11
3	O	261/304 (86%)	184 (70%)	54 (21%)	23 (9%)	1 11
4	D	151/186 (81%)	107 (71%)	26 (17%)	18 (12%)	0 6
4	P	150/186 (81%)	106 (71%)	26 (17%)	18 (12%)	0 6
5	E	212/214 (99%)	148 (70%)	43 (20%)	21 (10%)	0 9
5	Q	212/214 (99%)	148 (70%)	43 (20%)	21 (10%)	0 9
6	F	82/155 (53%)	57 (70%)	15 (18%)	10 (12%)	0 5
6	R	82/155 (53%)	57 (70%)	15 (18%)	10 (12%)	0 5
7	G	169/171 (99%)	129 (76%)	27 (16%)	13 (8%)	1 13
7	S	169/171 (99%)	130 (77%)	26 (15%)	13 (8%)	1 13
8	H	126/145 (87%)	94 (75%)	17 (14%)	15 (12%)	0 6
8	T	126/145 (87%)	94 (75%)	17 (14%)	15 (12%)	0 6
9	I	108/115 (94%)	70 (65%)	25 (23%)	13 (12%)	0 6
9	U	108/115 (94%)	69 (64%)	26 (24%)	13 (12%)	0 6
10	J	58/72 (81%)	34 (59%)	15 (26%)	9 (16%)	0 3

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
10	V	58/72 (81%)	34 (59%)	15 (26%)	9 (16%)	0 3
11	K	111/118 (94%)	82 (74%)	18 (16%)	11 (10%)	0 9
11	W	111/118 (94%)	82 (74%)	18 (16%)	11 (10%)	0 9
12	L	44/73 (60%)	19 (43%)	12 (27%)	13 (30%)	0 0
12	X	44/73 (60%)	19 (43%)	12 (27%)	13 (30%)	0 0
All	All	7446/9046 (82%)	5160 (69%)	1484 (20%)	802 (11%)	0 8

5 of 802 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	42	ASP
1	A	48	PRO
1	A	54	ASN
1	A	57	LYS
1	A	67	CYS

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	970/1528 (64%)	876 (90%)	94 (10%)	8 27
1	M	973/1528 (64%)	879 (90%)	94 (10%)	8 27
2	B	815/1077 (76%)	732 (90%)	83 (10%)	7 25
2	N	818/1077 (76%)	735 (90%)	83 (10%)	7 25
3	C	154/264 (58%)	133 (86%)	21 (14%)	3 17
3	O	157/264 (60%)	135 (86%)	22 (14%)	3 17
4	D	79/160 (49%)	68 (86%)	11 (14%)	3 17
4	P	80/160 (50%)	68 (85%)	12 (15%)	3 15
5	E	155/197 (79%)	141 (91%)	14 (9%)	9 30
5	Q	155/197 (79%)	141 (91%)	14 (9%)	9 30
6	F	60/137 (44%)	54 (90%)	6 (10%)	7 26

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
6	R	60/137 (44%)	54 (90%)	6 (10%)	7 26
7	G	102/148 (69%)	93 (91%)	9 (9%)	10 31
7	S	103/148 (70%)	94 (91%)	9 (9%)	10 31
8	H	90/130 (69%)	83 (92%)	7 (8%)	12 36
8	T	89/130 (68%)	82 (92%)	7 (8%)	12 35
9	I	80/109 (73%)	69 (86%)	11 (14%)	3 17
9	U	80/109 (73%)	69 (86%)	11 (14%)	3 17
10	J	47/66 (71%)	40 (85%)	7 (15%)	3 15
10	V	47/66 (71%)	40 (85%)	7 (15%)	3 15
11	K	68/109 (62%)	56 (82%)	12 (18%)	2 11
11	W	68/109 (62%)	56 (82%)	12 (18%)	2 11
12	L	25/58 (43%)	24 (96%)	1 (4%)	31 55
12	X	26/58 (45%)	25 (96%)	1 (4%)	33 57
All	All	5301/7966 (66%)	4747 (90%)	554 (10%)	7 24

5 of 554 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	O	145	CYS
4	P	41	HIS
3	O	136	ASP
8	T	94	TYR
3	C	233	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 137 such sidechains are listed below:

Mol	Chain	Res	Type
2	N	1193	GLN
4	P	126	GLN
8	T	138	ASN
4	D	99	ASN
3	C	167	HIS

5.3.3 RNA

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 16 ligands modelled in this entry, 16 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

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	1384/1743 (79%)	0.70	141 (10%) 6 10	201, 248, 299, 338	0
1	M	1386/1743 (79%)	0.82	177 (12%) 3 7	209, 256, 307, 345	0
2	B	1073/1227 (87%)	0.73	105 (9%) 7 10	200, 257, 309, 343	0
2	N	1074/1227 (87%)	0.81	130 (12%) 4 8	207, 264, 316, 350	0
3	C	265/304 (87%)	0.58	12 (4%) 33 31	211, 245, 280, 304	0
3	O	265/304 (87%)	0.73	33 (12%) 3 8	219, 253, 287, 312	0
4	D	163/186 (87%)	1.25	34 (20%) 1 3	229, 267, 306, 310	0
4	P	162/186 (87%)	1.03	26 (16%) 1 5	236, 275, 312, 318	0
5	E	214/214 (100%)	1.05	46 (21%) 0 3	228, 290, 334, 343	0
5	Q	214/214 (100%)	0.87	35 (16%) 1 5	236, 297, 342, 350	0
6	F	84/155 (54%)	0.66	5 (5%) 21 21	205, 231, 261, 268	0
6	R	84/155 (54%)	0.60	7 (8%) 11 13	213, 239, 268, 276	0
7	G	171/171 (100%)	0.87	21 (12%) 4 8	233, 252, 288, 298	0
7	S	171/171 (100%)	0.75	15 (8%) 10 12	240, 260, 295, 305	0
8	H	130/145 (89%)	1.04	24 (18%) 1 4	254, 286, 314, 330	0
8	T	130/145 (89%)	1.08	29 (22%) 0 2	261, 294, 322, 338	0
9	I	112/115 (97%)	1.30	27 (24%) 0 2	246, 286, 317, 326	0
9	U	112/115 (97%)	1.08	19 (16%) 1 4	254, 294, 325, 334	0
10	J	62/72 (86%)	0.26	1 (1%) 72 64	207, 236, 276, 290	0
10	V	62/72 (86%)	0.40	2 (3%) 47 41	214, 244, 283, 298	0
11	K	113/118 (95%)	0.82	13 (11%) 4 9	219, 246, 269, 294	0
11	W	113/118 (95%)	0.82	17 (15%) 2 5	226, 253, 276, 301	0
12	L	46/73 (63%)	1.76	13 (28%) 0 2	242, 315, 329, 332	0
12	X	46/73 (63%)	1.25	10 (21%) 0 3	249, 322, 337, 339	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
All	All	7636/9046 (84%)	0.81	942 (12%)	4	8	200, 258, 315, 350	0

The worst 5 of 942 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
12	L	28	GLY	10.3
12	X	28	GLY	8.6
3	O	268	ALA	8.6
2	B	724	LYS	7.9
12	X	27	TYR	7.6

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 monosaccharides 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
13	ZN	M	1801	1/1	0.20	0.20	249,249,249,249	0
13	ZN	X	101	1/1	0.39	0.17	278,278,278,278	0
13	ZN	U	201	1/1	0.41	0.13	266,266,266,266	0
13	ZN	I	201	1/1	0.62	0.37	258,258,258,258	0
13	ZN	L	101	1/1	0.71	0.13	270,270,270,270	0
13	ZN	A	1801	1/1	0.72	0.23	241,241,241,241	0
13	ZN	I	202	1/1	0.83	0.24	306,306,306,306	0
13	ZN	B	1301	1/1	0.85	0.11	228,228,228,228	0
13	ZN	O	401	1/1	0.86	0.09	238,238,238,238	0
13	ZN	C	401	1/1	0.90	0.07	230,230,230,230	0
13	ZN	A	1802	1/1	0.90	0.15	216,216,216,216	0
13	ZN	U	202	1/1	0.93	0.07	314,314,314,314	0
13	ZN	M	1802	1/1	0.93	0.19	224,224,224,224	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
13	ZN	J	101	1/1	0.96	0.31	234,234,234,234	0
13	ZN	V	101	1/1	0.97	0.27	242,242,242,242	0
13	ZN	N	1301	1/1	0.98	0.15	236,236,236,236	0

6.5 Other polymers (i)

There are no such residues in this entry.