



wwPDB EM Map/Model Validation Report ⓘ

Sep 27, 2016 – 02:27 PM EDT

PDB ID : 5SVA
EMDB ID: : EMD-8305
Title : Mediator-RNA Polymerase II Pre-Initiation Complex
Authors : Robinson, P.J.; Bushnell, D.A.; Kornberg, R.D.
Deposited on : 2016-08-05
Resolution : 15.30 Å(reported)

This is a wwPDB EM Map/Model Validation Report for a publicly released PDB/EMDB entry.
For rigid body fitted models, validation errors reported here could stem from errors in the original structure(s) used in the fitting.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
<http://wwpdb.org/validation/2016/EMValidationReportHelp>

MolProbity : 4.02b-467
Mogul : unknown
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20027939

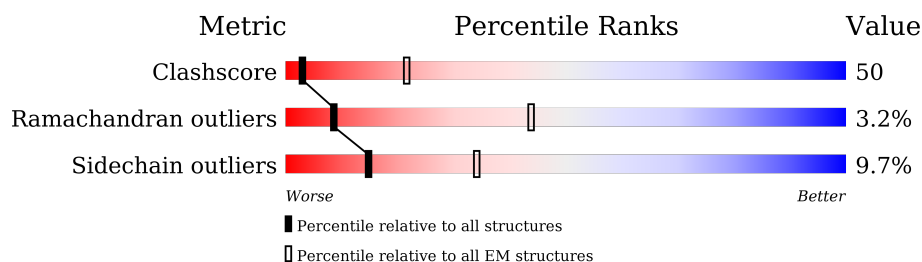
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 15.30 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)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

The table below summarises the geometric issues observed across the polymeric chains. The red, orange, yellow and green segments on the 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 $\leq 5\%$

Mol	Chain	Length	Quality of chain
1	A	1733	
2	B	1224	
3	C	318	
4	D	221	
5	E	215	
6	F	155	
7	G	171	
8	H	146	
9	I	122	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
10	J	70	
11	K	120	
12	L	70	
13	M	295	
14	N	223	
15	O	115	
16	P	687	
17	Q	307	
18	R	210	
19	S	121	
20	T	284	
21	U	222	
22	V	149	
23	W	140	
24	X	127	
25	Y	778	
26	Z	843	
27	a	513	
28	b	72	
29	c	345	
30	d	286	
31	e	122	
32	f	735	
33	g	400	
34	h	482	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
35	i	328	<div><div></div><div>28%8%63%</div></div>
36	j	240	<div><div></div><div>70%25%</div></div>
37	k	25	<div><div></div><div>68%32%</div></div>
38	l	108	<div><div></div><div>57%43%</div></div>
39	m	108	<div><div></div><div>57%43%</div></div>
40	n	244	<div><div></div><div>82%18%</div></div>

2 Entry composition

There are 42 unique types of molecules in this entry. The entry contains 66759 atoms, of which 626 are hydrogens and 0 are deuteriums.

In the tables below, 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					AltConf	Trace
1	A	1422	Total	C	N	O	S	0	0
			11174	7036	1954	2122	62		

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

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	1156	Total	C	N	O	S	0	0
			9140	5781	1606	1697	56		

- Molecule 3 is a protein called DNA-directed RNA polymerase II subunit RPB3.

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

- Molecule 4 is a protein called DNA-directed RNA polymerase II subunit RPB4.

Mol	Chain	Residues	Atoms					AltConf	Trace
4	D	178	Total	C	N	O	S	0	0
			1434	887	257	288	2		

- Molecule 5 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC1.

Mol	Chain	Residues	Atoms					AltConf	Trace
5	E	214	Total	C	N	O	S	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					AltConf	Trace
6	F	84	Total	C	N	O	S	0	0
			679	434	115	127	3		

- Molecule 7 is a protein called DNA-directed RNA polymerase II subunit RPB7.

Mol	Chain	Residues	Atoms					AltConf	Trace
7	G	171	Total	C	N	O	S	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					AltConf	Trace
8	H	133	Total	C	N	O	S	0	0
			1068	673	180	211	4		

- Molecule 9 is a protein called DNA-directed RNA polymerase II subunit RPB9.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	119	Total	C	N	O	S	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					AltConf	Trace
10	J	65	Total	C	N	O	S	0	0
			532	339	93	94	6		

- Molecule 11 is a protein called DNA-directed RNA polymerase II subunit RPB11.

Mol	Chain	Residues	Atoms					AltConf	Trace
11	K	115	Total	C	N	O	S	0	1
			920	590	157	171	2		

- Molecule 12 is a protein called DNA-directed RNA polymerases I, II, and III subunit RPABC4.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	46	Total	C	N	O	S	0	0
			363	224	72	63	4		

- Molecule 13 is a protein called Mediator of RNA polymerase II transcription subunit 6.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	156	Total	C	N	O	S	0	0
			777	464	156	156	1		

- Molecule 14 is a protein called Mediator of RNA polymerase II transcription subunit 8.

Mol	Chain	Residues	Atoms				AltConf	Trace
14	N	168	Total	C	N	O	0	0
			891	542	172	177		

- Molecule 15 is a protein called Mediator of RNA polymerase II transcription subunit 11.

Mol	Chain	Residues	Atoms				AltConf	Trace
15	O	103	Total	C	N	O	0	0
			511	305	103	103		

- Molecule 16 is a protein called Mediator of RNA polymerase II transcription subunit 17.

Mol	Chain	Residues	Atoms				AltConf	Trace
16	P	487	Total	C	N	O	0	0
			2421	1447	487	487		

- Molecule 17 is a protein called Mediator of RNA polymerase II transcription subunit 18.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	Q	253	Total	C	N	O	S	0	0
			1979	1255	330	384	10		

- Molecule 18 is a protein called Mediator of RNA polymerase II transcription subunit 20.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	R	209	Total	C	N	O	S	0	0
			1600	1011	269	315	5		

- Molecule 19 is a protein called Mediator of RNA polymerase II transcription subunit 22.

Mol	Chain	Residues	Atoms				AltConf	Trace
19	S	109	Total	C	N	O	0	0
			544	326	109	109		

- Molecule 20 is a protein called Mediator of RNA polymerase II transcription subunit 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	T	91	Total	C	N	O	S	0	0
			756	475	125	154	2		

- Molecule 21 is a protein called Mediator of RNA polymerase II transcription subunit 7.

Mol	Chain	Residues	Atoms					AltConf	Trace
21	U	156	Total	C	N	O	S	0	0
			1310	847	220	238	5		

- Molecule 22 is a protein called Mediator of RNA polymerase II transcription subunit 9.

Mol	Chain	Residues	Atoms					AltConf	Trace
22	V	85	Total	C	N	O	S	0	0
			720	451	133	135	1		

- Molecule 23 is a protein called Mediator of RNA polymerase II transcription subunit 21.

Mol	Chain	Residues	Atoms					AltConf	Trace
23	W	119	Total	C	N	O	S	0	0
			965	608	160	193	4		

- Molecule 24 is a protein called Mediator of RNA polymerase II transcription subunit 31.

Mol	Chain	Residues	Atoms					AltConf	Trace
24	X	92	Total	C	N	O	S	0	0
			767	506	116	141	4		

- Molecule 25 is a protein called DNA repair helicase RAD3.

Mol	Chain	Residues	Atoms						AltConf	Trace
25	Y	562	Total	C	H	N	O	S	0	0
			5175	2901	626	777	838	33		

- Molecule 26 is a protein called DNA repair helicase RAD25.

Mol	Chain	Residues	Atoms					AltConf	Trace
26	Z	469	Total	C	N	O	S	0	0
			3769	2370	660	716	23		

- Molecule 27 is a protein called RNA polymerase II transcription factor B subunit 2.

Mol	Chain	Residues	Atoms				AltConf	Trace
27	a	62	Total	C	N	O	0	0
			518	334	83	101		

- Molecule 28 is a protein called RNA polymerase II transcription factor B subunit 5.

Mol	Chain	Residues	Atoms					AltConf	Trace
28	b	63	Total	C	N	O	S	0	0
			499	316	88	93	2		

- Molecule 29 is a protein called Transcription initiation factor IIB.

Mol	Chain	Residues	Atoms					AltConf	Trace
29	c	189	Total	C	N	O	S	0	0
			1357	838	240	267	12		

- Molecule 30 is a protein called Transcription initiation factor IIA large subunit.

Mol	Chain	Residues	Atoms					AltConf	Trace
30	d	116	Total	C	N	O	S	0	0
			956	599	159	195	3		

- Molecule 31 is a protein called Transcription initiation factor IIA subunit 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
31	e	101	Total	C	N	O	S	0	0
			792	500	132	156	4		

- Molecule 32 is a protein called Transcription initiation factor IIF subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
32	f	149	Total	C	N	O	S	0	0
			1243	788	222	229	4		

- Molecule 33 is a protein called Transcription initiation factor IIF subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
33	g	172	Total	C	N	O	S	0	0
			1443	922	248	267	6		

- Molecule 34 is a protein called Transcription initiation factor IIE subunit alpha.

Mol	Chain	Residues	Atoms					AltConf	Trace
34	h	118	Total	C	N	O	S	0	0
			960	625	158	172	5		

- Molecule 35 is a protein called Transcription initiation factor IIE subunit beta.

Mol	Chain	Residues	Atoms					AltConf	Trace
35	i	120	Total	C	N	O	S	0	0
			987	636	161	187	3		

- Molecule 36 is a protein called TATA-box-binding protein.

Mol	Chain	Residues	Atoms					AltConf	Trace
36	j	180	Total	C	N	O	S	0	0
			1416	921	242	247	6		

- Molecule 37 is a protein called DNA-directed RNA polymerase II subunit RPB1.

Mol	Chain	Residues	Atoms				AltConf	Trace
37	k	25	Total	C	N	O	0	0
			184	116	25	43		

- Molecule 38 is a DNA chain called 108bp HIS4 Promoter Non-template Strand (-92/+16).

Mol	Chain	Residues	Atoms					AltConf	Trace
38	l	62	Total	C	N	O	P	0	0
			1271	609	222	378	62		

- Molecule 39 is a DNA chain called 108bp HIS4 Promoter Template Strand (+16/-92).

Mol	Chain	Residues	Atoms					AltConf	Trace
39	m	62	Total	C	N	O	P	0	0
			1271	607	236	366	62		

- Molecule 40 is a protein called Transcription initiation factor TFIID subunit 14.

Mol	Chain	Residues	Atoms		AltConf	Trace
40	n	200	Total	C	0	200
			200	200		

- Molecule 41 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
41	J	1	Total	Zn	0
			1	1	
41	B	1	Total	Zn	0
			1	1	
41	I	2	Total	Zn	0
			2	2	

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		AltConf
41	C	1	Total 1	Zn 1	0
41	A	2	Total 2	Zn 2	0
41	L	1	Total 1	Zn 1	0

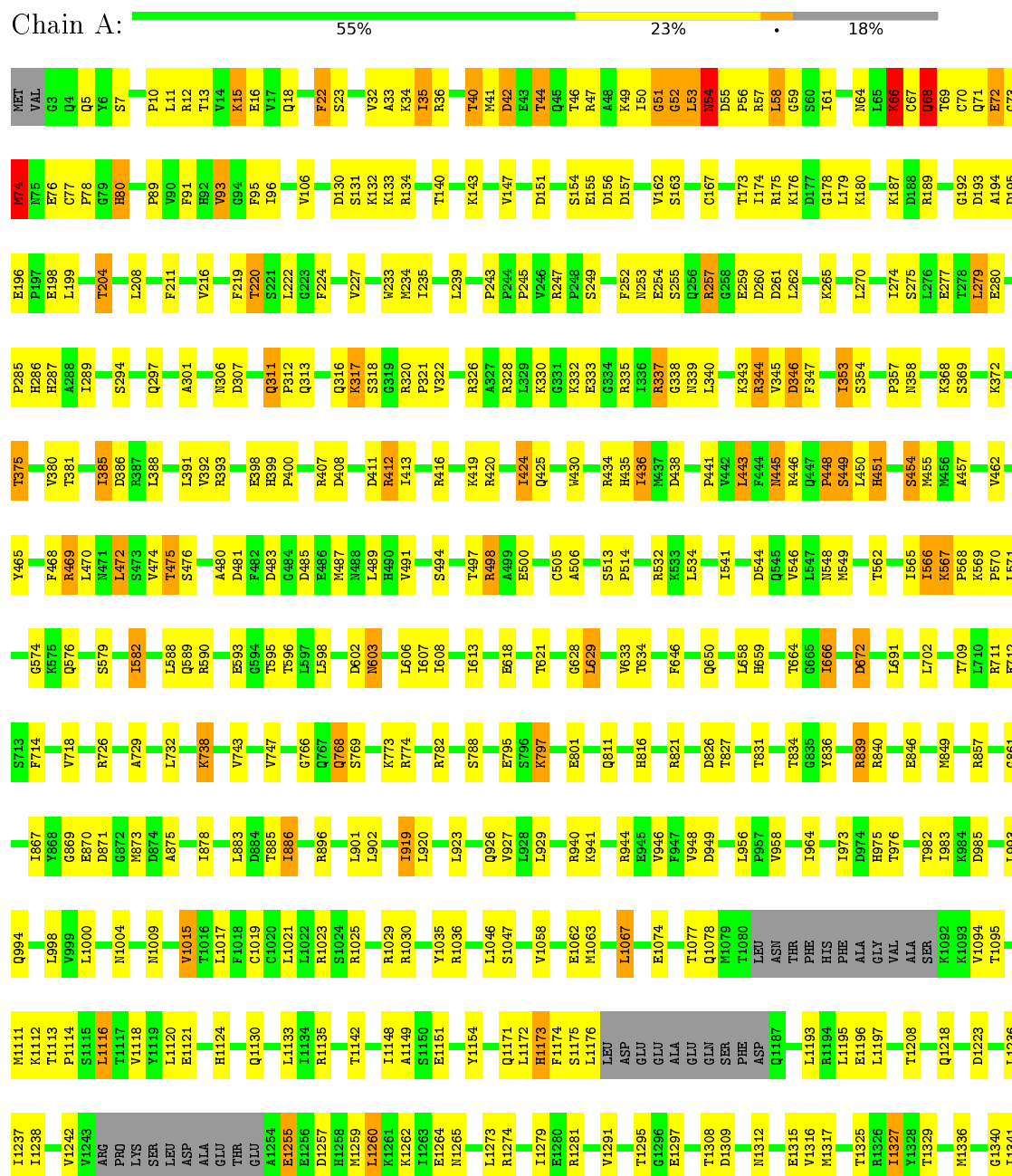
- Molecule 42 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

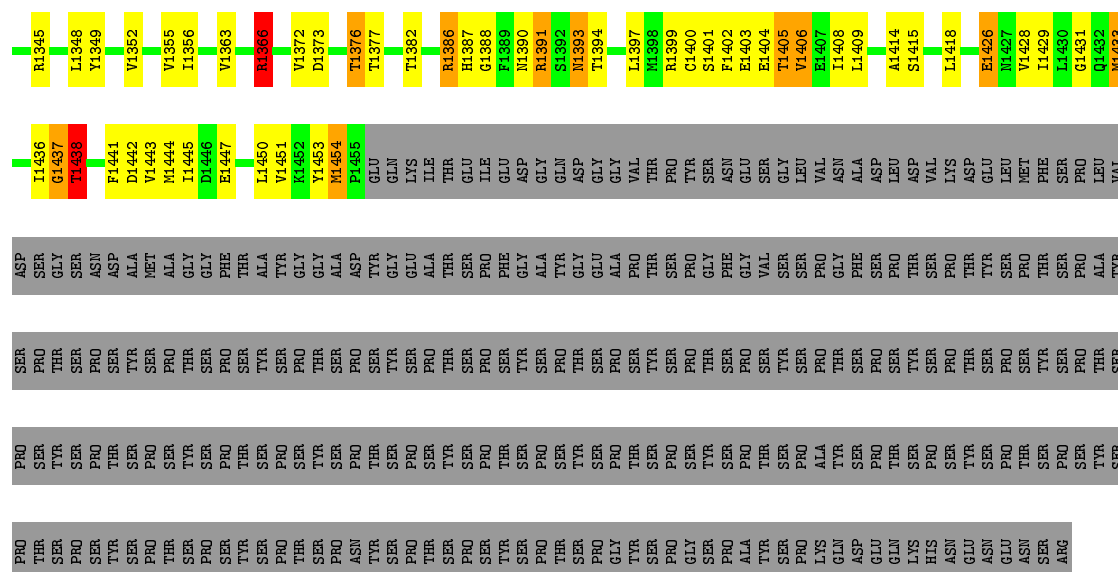
Mol	Chain	Residues	Atoms		AltConf
42	A	1	Total 1	Mg 1	0

3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

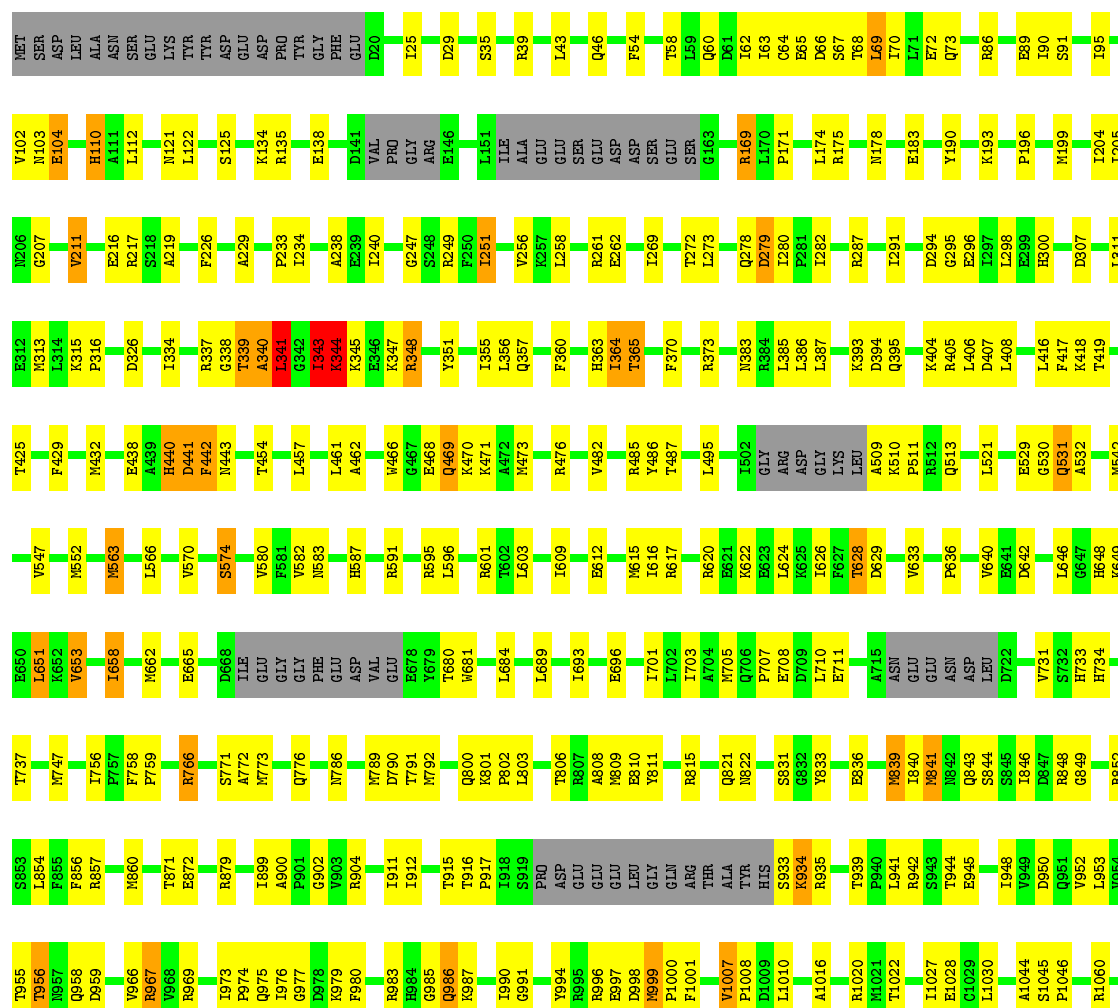
- Molecule 1: DNA-directed RNA polymerase II subunit RPB1

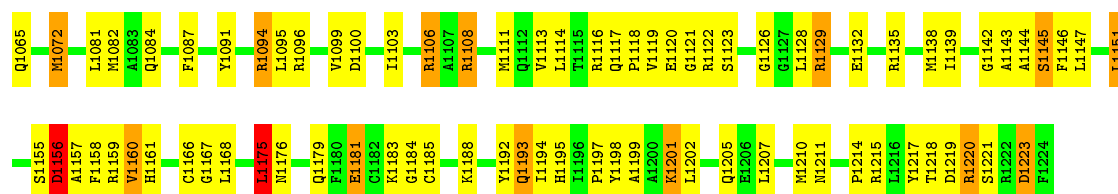




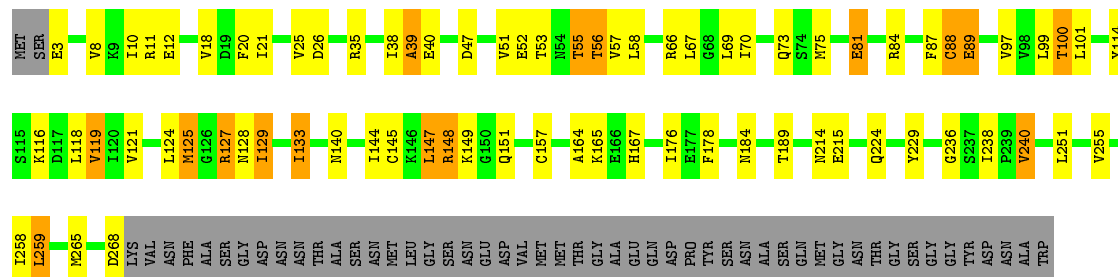
• Molecule 2: DNA-directed RNA polymerase II subunit RPB2

Chain B: 62% 28% 6%

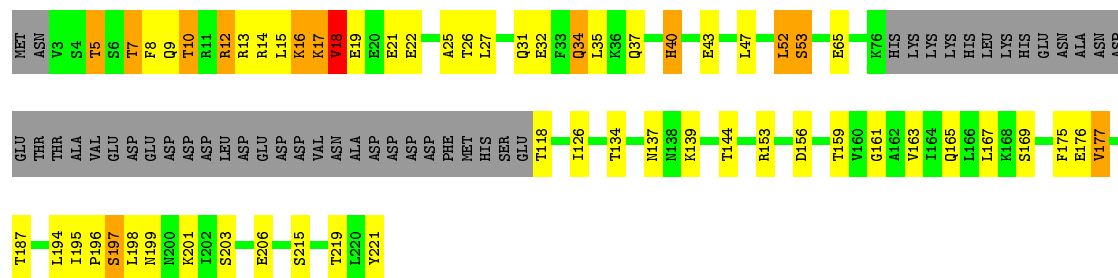




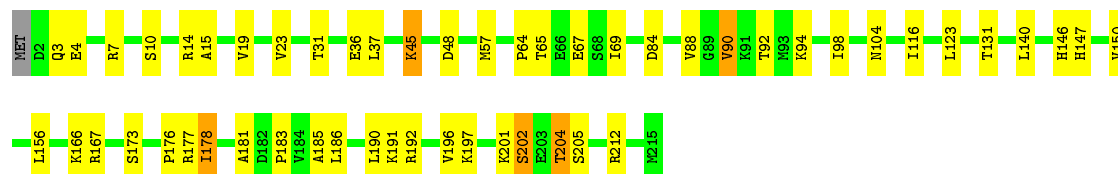
• Molecule 3: DNA-directed RNA polymerase II subunit RPB3



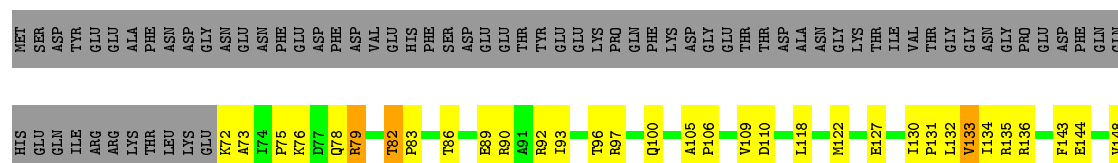
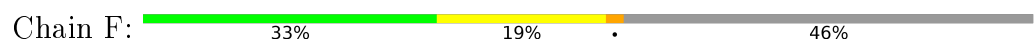
• Molecule 4: DNA-directed RNA polymerase II subunit RPB4



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



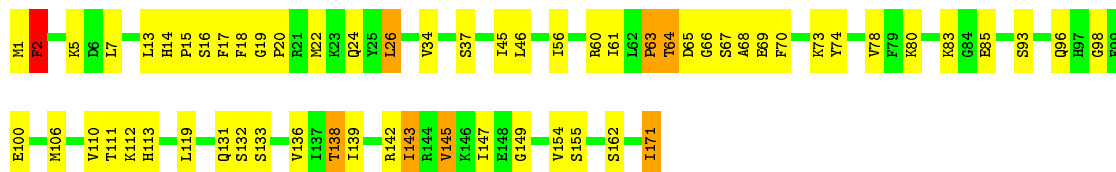
• Molecule 6: DNA-directed RNA polymerases I, II, and III subunit RPABC2





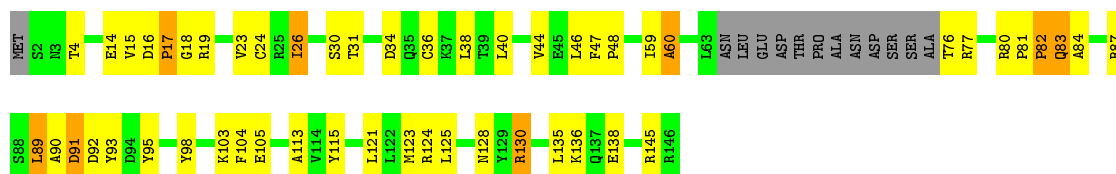
- Molecule 7: DNA-directed RNA polymerase II subunit RPB7

Chain G: 64% 31%



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

Chain H: 55% 30% 5% 9%



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

Chain I: 75% 20%



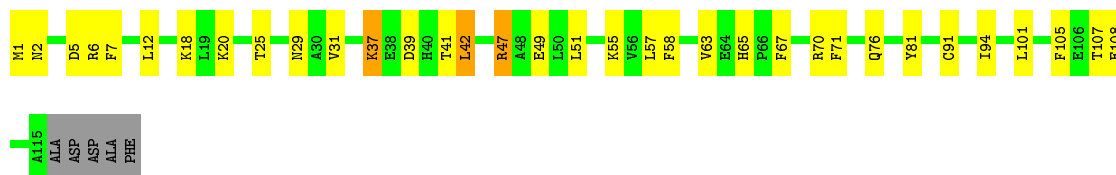
- Molecule 10: DNA-directed RNA polymerases I, II, and III subunit RPABC5

Chain J: 56% 26% 11% 7%



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

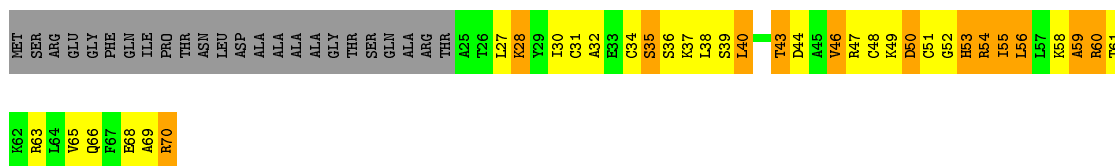
Chain K: 68% 26%



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

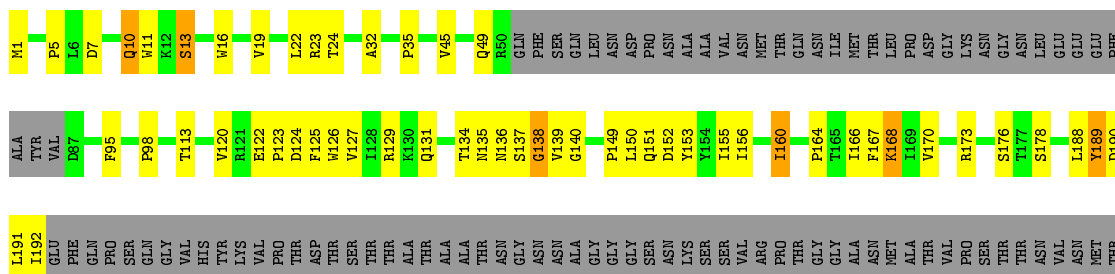
Chain L: 16% 31% 19% 34%





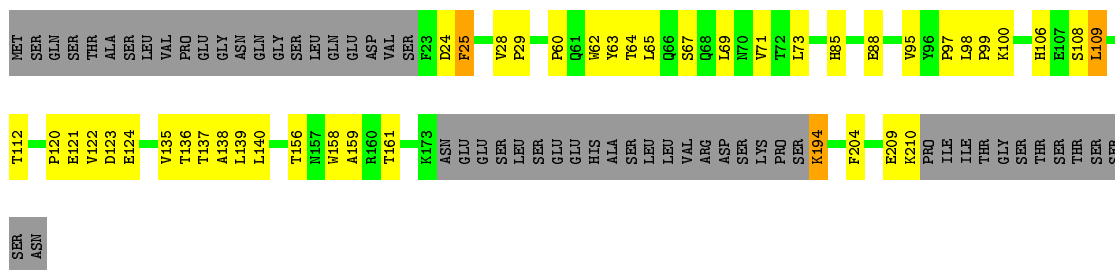
- Molecule 13: Mediator of RNA polymerase II transcription subunit 6

Chain M: 34% 17% 47%



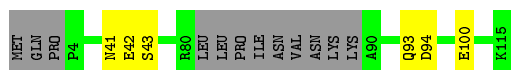
- Molecule 14: Mediator of RNA polymerase II transcription subunit 8

Chain N: 56% 18% 25%



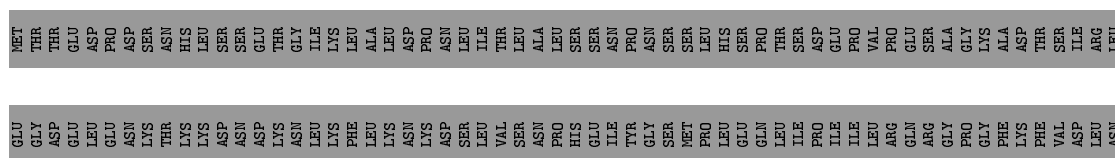
- Molecule 15: Mediator of RNA polymerase II transcription subunit 11

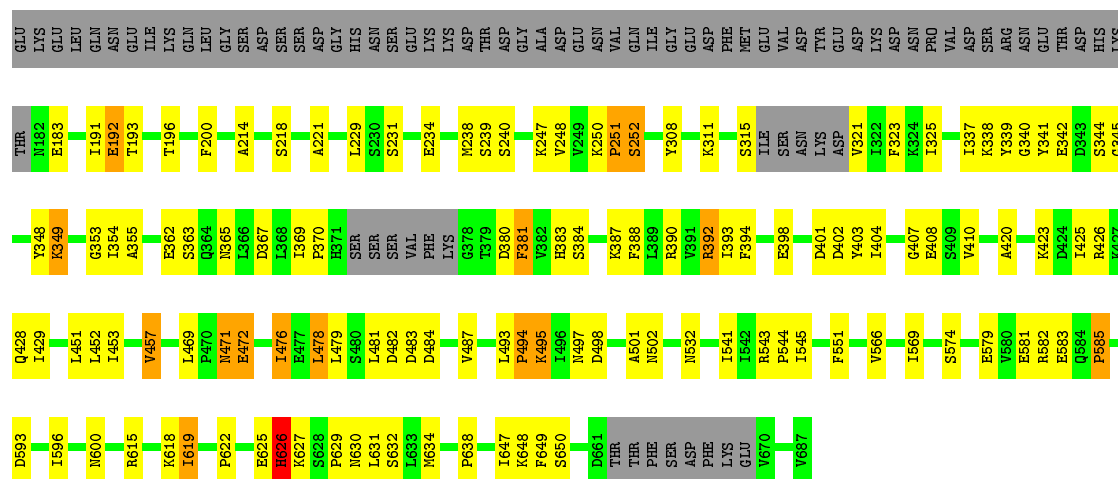
Chain O: 84% 5% 10%



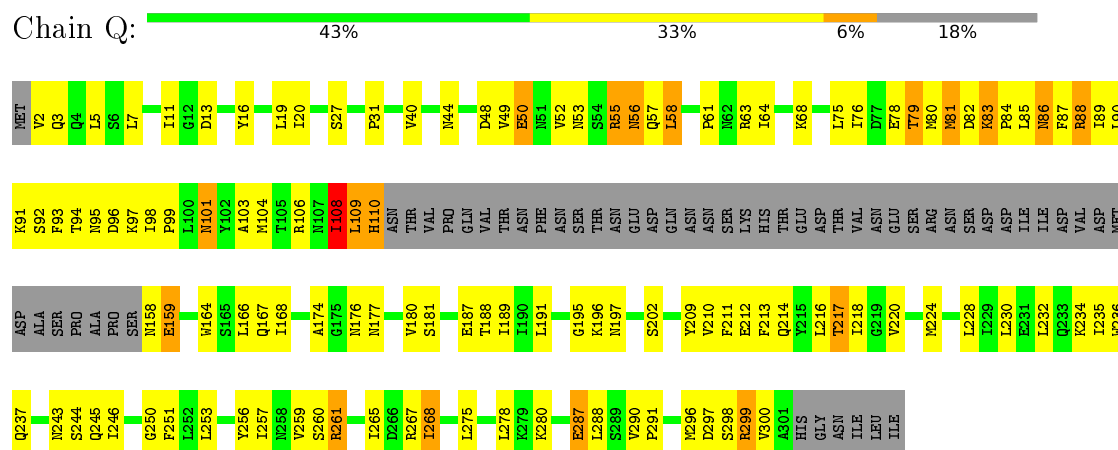
- Molecule 16: Mediator of RNA polymerase II transcription subunit 17

Chain P: 53% 16% 29%

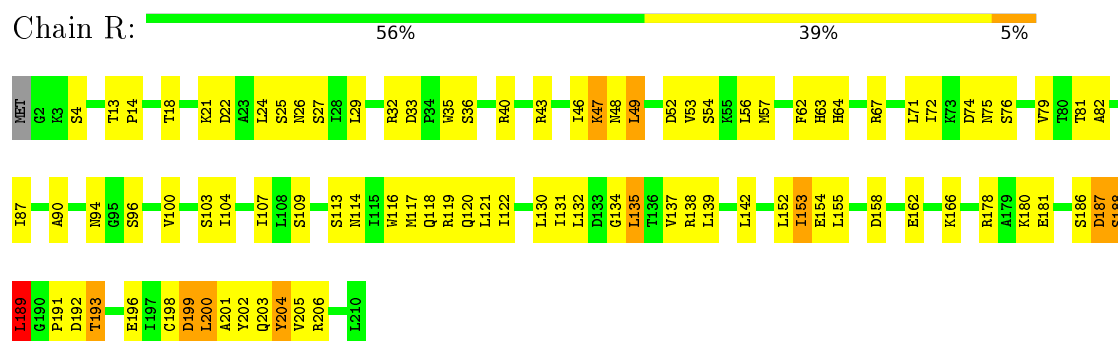




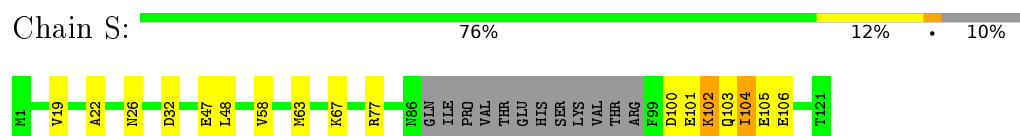
- Molecule 17: Mediator of RNA polymerase II transcription subunit 18



- Molecule 18: Mediator of RNA polymerase II transcription subunit 20



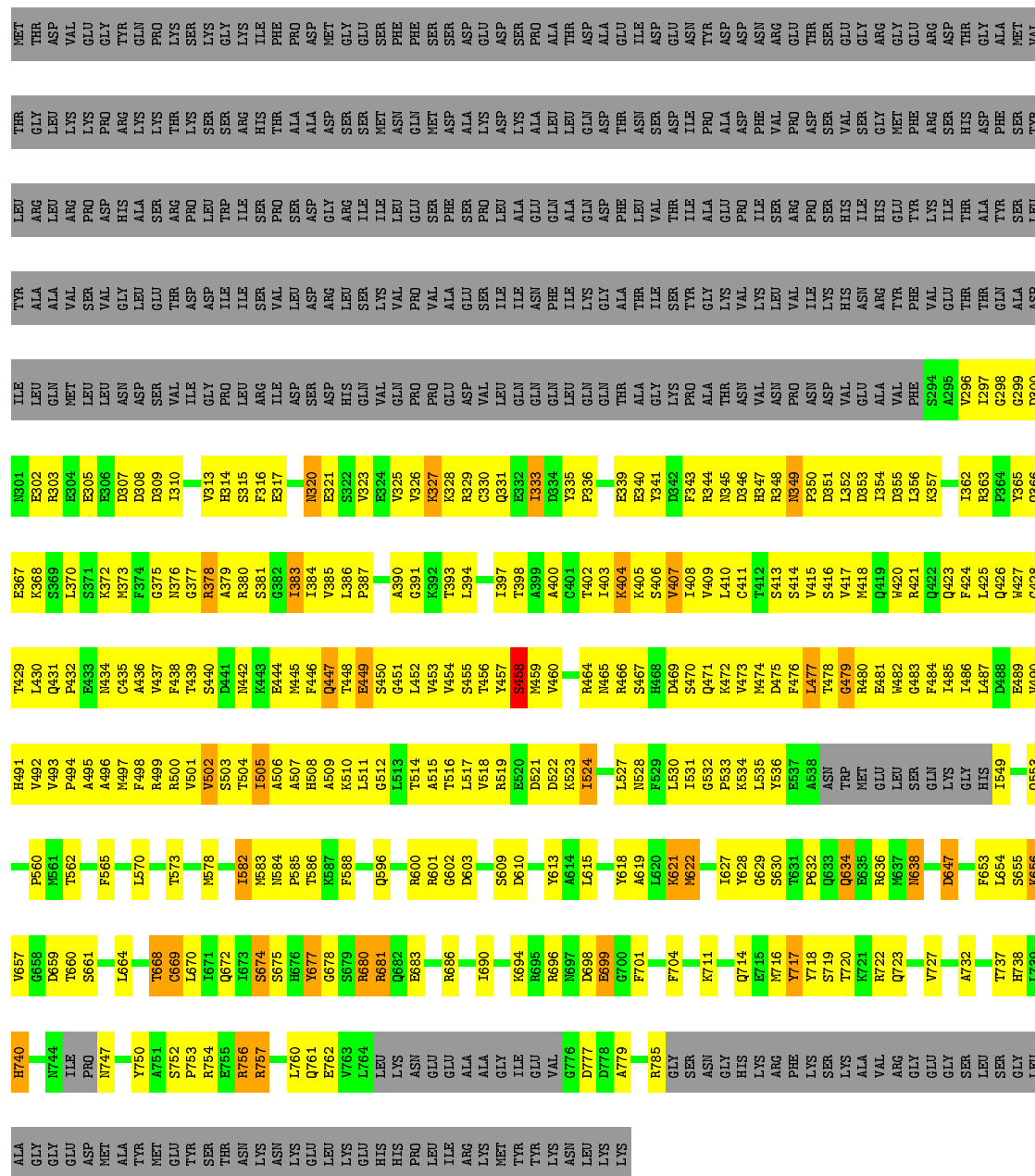
- Molecule 19: Mediator of RNA polymerase II transcription subunit 22



- Molecule 20: Mediator of RNA polymerase II transcription subunit 4

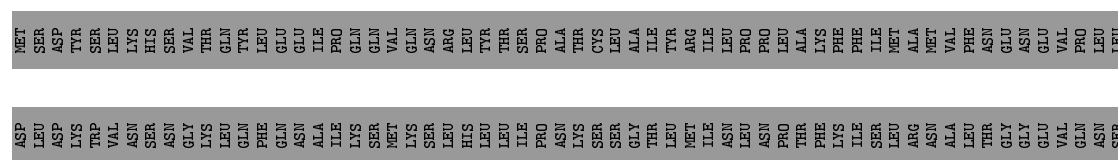
- Molecule 26: DNA repair helicase RAD25

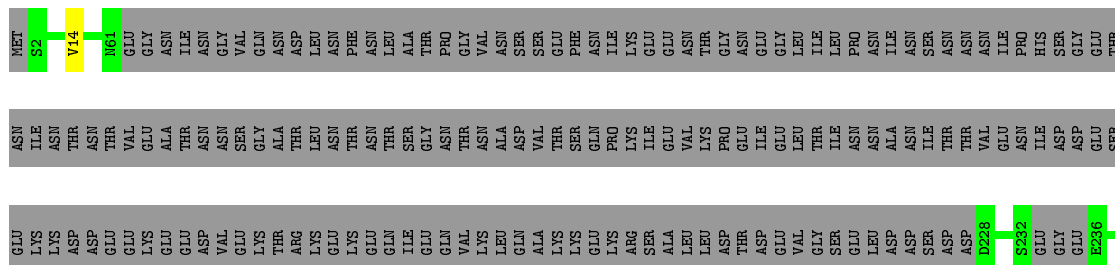
Chain Z:  22% 30% . 44%



- Molecule 27: RNA polymerase II transcription factor B subunit 2

Chain a: 12% 88%







ARG
LYS
GLY
LYS
ILE
THR
ASN
THR
HIS
MET
THR
GLY
ILE
LEU
LYS
ASP
TYR
SER
HIS
ARG
VAL

- Molecule 36: TATA-box-binding protein

Chain j:  70% 5% 25%

MET	ALA	ALA	ASP	GLU	GLU	ARG	LEU	LYS	PHE	GLY	LYS	GLU	ALA	ASN	LYS	ILE	VAL	PHE	ASP	PRO	ASN	THR	ARG	GIN	VAL	TRP	GLU	ASN	GLN	ASN	ARG	ASP	GLY	LYS	PRO	ALA	THR	THR	PHE	GLN	SER	GLY	GLU	GLY	ASP	ILE	LYS	ANG	ALA	ALA	PRO	GLU	SER	THR	ASP	LYS	ALA
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Category	Count
S61	10
Q68	10
C78	10
L87	10
K110	10
R141	10
K145	10
K151	10
D154	10
K211	10
V233	10
L234	10
M240	10

- Molecule 37: DNA-directed RNA polymerase II subunit RPB1

Chain k:  68% 32%

P1
S4
S9 Y10
S14 P15
P19 T20 S21
S25

- Molecule 38: 108bp HIS4 Promoter Non-template Strand (-92/+16)

Chain 1: 57% 43%

[illegible]

- Molecule 39: 108bp HIS4 Promoter Template Strand (+16/-92)

Chain m:  57% 43%

[illegible]

- Molecule 40: Transcription initiation factor TFIID subunit 14

Chain n: 82% 18%

M1	T137	GLU	GLU	THR	THR	ASN	THR	GLY	ILE	GLY	LNS	ARG	ARG	THR	THR	ASN	THR	THR	ALA	GLU	PRO	LNS	ALA	LNS	ARG	ALA	LNS	GLY	SER	ALA	SER	SER	THR	VAL	LNS	GLY	SER	VAL	D178	E240	LNS	ASN	THR	VAL
----	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	-----	-----	-----	-----

4 Experimental information

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	170600	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	Not provided	Depositor
Microscope	FEI TECNAI 20	Depositor
Voltage (kV)	200	Depositor
Electron dose ($e^-/\text{\AA}^2$)	Not provided	Depositor
Minimum defocus (nm)	Not provided	Depositor
Maximum defocus (nm)	Not provided	Depositor
Magnification	Not provided	Depositor
Image detector	Not provided	Depositor

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z > 2$	RMSZ	$\# Z > 2$
1	A	0.51	1/11374 (0.0%)	0.79	5/15384 (0.0%)
10	J	0.54	0/541	0.88	1/727 (0.1%)
11	K	0.48	0/938	0.71	0/1267
12	L	0.54	0/365	0.79	0/485
13	M	0.61	0/775	0.83	0/1077
14	N	0.53	0/893	0.76	0/1237
15	O	0.52	0/509	0.67	0/707
16	P	0.58	0/2417	0.79	2/3369 (0.1%)
17	Q	0.56	0/2014	0.75	0/2728
18	R	0.50	2/1626 (0.1%)	0.66	0/2205
19	S	0.57	0/542	0.73	0/755
2	B	0.48	0/9317	0.74	4/12567 (0.0%)
20	T	0.69	0/763	1.10	2/1025 (0.2%)
21	U	0.43	0/1339	0.60	0/1808
22	V	0.73	0/732	1.01	4/984 (0.4%)
23	W	0.47	0/973	0.64	0/1308
24	X	0.39	0/789	0.53	0/1077
25	Y	0.55	2/4616 (0.0%)	0.79	13/6196 (0.2%)
26	Z	0.78	0/3837	0.98	8/5177 (0.2%)
27	a	0.67	0/527	0.68	0/704
28	b	0.60	0/504	0.69	1/679 (0.1%)
29	c	0.29	0/1373	0.47	0/1863
3	C	0.49	0/2133	0.78	2/2891 (0.1%)
30	d	0.40	0/970	0.57	0/1310
31	e	0.42	0/800	0.63	0/1080
32	f	0.33	0/1267	0.82	9/1700 (0.5%)
33	g	0.66	0/1469	0.73	3/1972 (0.2%)
34	h	0.96	0/978	1.11	1/1321 (0.1%)
35	i	0.37	0/1003	0.61	0/1345
36	j	0.41	0/1443	0.62	0/1942
37	k	0.77	0/194	0.69	0/270
38	l	0.43	0/1423	0.89	0/2195

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	# Z >2	RMSZ	# Z >2
39	m	0.47	0/1427	0.89	0/2199
4	D	0.51	0/1444	0.83	2/1935 (0.1%)
5	E	0.48	0/1788	0.72	0/2406
6	F	0.58	0/691	0.79	0/933
7	G	0.51	0/1368	0.81	0/1844
8	H	0.51	0/1086	0.80	0/1470
9	I	0.47	0/989	0.78	0/1331
All	All	0.54	5/67237 (0.0%)	0.78	57/91473 (0.1%)

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
13	M	0	2
16	P	0	3
17	Q	0	1
2	B	0	1
25	Y	0	1
26	Z	0	1
All	All	0	9

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1394	THR	C-N	-6.29	1.21	1.33
25	Y	172	PRO	N-CD	5.61	1.55	1.47
25	Y	19	PRO	N-CD	5.51	1.55	1.47
18	R	35	TRP	CD2-CE2	5.06	1.47	1.41
18	R	116	TRP	CD2-CE2	5.04	1.47	1.41

The worst 5 of 57 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
25	Y	493	LEU	C-N-CD	-13.71	90.45	120.60
25	Y	533	THR	C-N-CD	-13.65	90.57	120.60
32	f	135	LEU	C-N-CD	-11.09	96.21	120.60
26	Z	505	ILE	CB-CA-C	-9.99	91.62	111.60
25	Y	433	PRO	CA-N-CD	-9.03	98.86	111.50

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	43	LEU	Mainchain
13	M	10	GLN	Peptide
13	M	138	GLY	Peptide
16	P	471	ASN	Peptide
16	P	532	ASN	Peptide

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	11174	0	11225	498	0
2	B	9140	0	9111	297	0
3	C	2095	0	2051	62	0
4	D	1434	0	1460	70	0
5	E	1752	0	1776	26	0
6	F	679	0	698	109	0
7	G	1340	0	1355	228	0
8	H	1068	0	1040	25	0
9	I	971	0	927	15	0
10	J	532	0	542	14	0
11	K	920	0	929	21	0
12	L	363	0	386	54	0
13	M	777	0	329	39	0
14	N	891	0	455	25	0
15	O	511	0	215	3	0
16	P	2421	0	1017	101	0
17	Q	1979	0	1977	245	0
18	R	1600	0	1614	107	0
19	S	544	0	241	10	0
20	T	756	0	761	265	0
21	U	1310	0	1339	106	0
22	V	720	0	725	280	0
23	W	965	0	987	136	0
24	X	767	0	752	22	0
25	Y	4549	626	4642	1289	0
26	Z	3769	0	3697	845	0
27	a	518	0	514	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
28	b	499	0	525	0	0
29	c	1357	0	1266	0	0
30	d	956	0	916	0	0
31	e	792	0	806	0	0
32	f	1243	0	1238	0	0
33	g	1443	0	1461	0	0
34	h	960	0	973	0	0
35	i	987	0	999	0	0
36	j	1416	0	1491	0	0
37	k	184	0	163	0	0
38	l	1271	0	705	0	0
39	m	1271	0	699	0	0
40	n	200	0	0	0	0
41	A	2	0	0	0	0
41	B	1	0	0	0	0
41	C	1	0	0	0	0
41	I	2	0	0	0	0
41	J	1	0	0	0	0
41	L	1	0	0	0	0
42	A	1	0	0	0	0
All	All	66133	626	62007	4244	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 50.

The worst 5 of 4244 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)
25:Y:254:THR:HG23	25:Y:286:TYR:CE1	1.20	1.71
25:Y:352:ILE:HB	25:Y:380:ARG:CZ	1.23	1.66
26:Z:757:ARG:HH12	26:Z:760:LEU:CG	1.04	1.66
25:Y:550:ILE:HG22	25:Y:554:TRP:CZ2	1.17	1.63
25:Y:639:LEU:HD12	25:Y:653:PHE:CE2	1.32	1.61

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

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	1414/1733 (82%)	1254 (89%)	112 (8%)	48 (3%)	5	40
2	B	1142/1224 (93%)	1022 (90%)	84 (7%)	36 (3%)	5	41
3	C	264/318 (83%)	242 (92%)	20 (8%)	2 (1%)	24	69
4	D	174/221 (79%)	149 (86%)	17 (10%)	8 (5%)	3	33
5	E	212/215 (99%)	195 (92%)	13 (6%)	4 (2%)	10	52
6	F	82/155 (53%)	75 (92%)	7 (8%)	0	100	100
7	G	169/171 (99%)	157 (93%)	9 (5%)	3 (2%)	11	53
8	H	129/146 (88%)	106 (82%)	14 (11%)	9 (7%)	1	22
9	I	117/122 (96%)	98 (84%)	16 (14%)	3 (3%)	7	45
10	J	63/70 (90%)	51 (81%)	9 (14%)	3 (5%)	3	32
11	K	113/120 (94%)	109 (96%)	4 (4%)	0	100	100
12	L	44/70 (63%)	19 (43%)	14 (32%)	11 (25%)	0	2
13	M	152/295 (52%)	115 (76%)	20 (13%)	17 (11%)	0	11
14	N	164/223 (74%)	133 (81%)	16 (10%)	15 (9%)	1	17
15	O	99/115 (86%)	92 (93%)	4 (4%)	3 (3%)	5	42
16	P	479/687 (70%)	385 (80%)	61 (13%)	33 (7%)	1	23
17	Q	249/307 (81%)	217 (87%)	25 (10%)	7 (3%)	6	44
18	R	207/210 (99%)	190 (92%)	12 (6%)	5 (2%)	7	47
19	S	105/121 (87%)	91 (87%)	7 (7%)	7 (7%)	1	24
20	T	89/284 (31%)	69 (78%)	20 (22%)	0	100	100
21	U	150/222 (68%)	127 (85%)	22 (15%)	1 (1%)	26	71
22	V	83/149 (56%)	73 (88%)	5 (6%)	5 (6%)	2	26
23	W	115/140 (82%)	95 (83%)	18 (16%)	2 (2%)	11	55
24	X	90/127 (71%)	86 (96%)	4 (4%)	0	100	100
25	Y	534/778 (69%)	503 (94%)	23 (4%)	8 (2%)	13	57

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
26	Z	461/843 (55%)	430 (93%)	26 (6%)	5 (1%)	17	63
27	a	60/513 (12%)	60 (100%)	0	0	100	100
28	b	61/72 (85%)	58 (95%)	3 (5%)	0	100	100
29	c	185/345 (54%)	164 (89%)	19 (10%)	2 (1%)	17	63
30	d	110/286 (38%)	103 (94%)	7 (6%)	0	100	100
31	e	97/122 (80%)	93 (96%)	4 (4%)	0	100	100
32	f	143/735 (20%)	130 (91%)	9 (6%)	4 (3%)	6	44
33	g	164/400 (41%)	148 (90%)	12 (7%)	4 (2%)	7	47
34	h	112/482 (23%)	100 (89%)	10 (9%)	2 (2%)	11	53
35	i	114/328 (35%)	102 (90%)	9 (8%)	3 (3%)	7	45
36	j	178/240 (74%)	170 (96%)	5 (3%)	3 (2%)	11	55
37	k	23/25 (92%)	9 (39%)	6 (26%)	8 (35%)	0	0
All	All	8147/12614 (65%)	7220 (89%)	666 (8%)	261 (3%)	8	41

5 of 261 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	74	MET
1	A	189	ARG
1	A	195	ASP
1	A	286	HIS
1	A	317	LYS

5.3.2 Protein sidechains ⓘ

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all PDB entries followed by that with respect to all EM entries.

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	1240/1520 (82%)	1071 (86%)	169 (14%)	5	27
2	B	985/1061 (93%)	868 (88%)	117 (12%)	6	31
3	C	234/274 (85%)	206 (88%)	28 (12%)	6	31
4	D	160/200 (80%)	129 (81%)	31 (19%)	2	12

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
5	E	196/197 (100%)	175 (89%)	21 (11%)	8	36
6	F	74/137 (54%)	67 (90%)	7 (10%)	11	41
7	G	152/152 (100%)	137 (90%)	15 (10%)	10	39
8	H	117/128 (91%)	103 (88%)	14 (12%)	6	31
9	I	113/116 (97%)	106 (94%)	7 (6%)	23	60
10	J	60/65 (92%)	49 (82%)	11 (18%)	2	14
11	K	99/102 (97%)	87 (88%)	12 (12%)	6	31
12	L	40/57 (70%)	37 (92%)	3 (8%)	17	53
13	M	1/259 (0%)	1 (100%)	0	100	100
14	N	16/207 (8%)	14 (88%)	2 (12%)	6	30
17	Q	226/280 (81%)	206 (91%)	20 (9%)	12	45
18	R	177/178 (99%)	162 (92%)	15 (8%)	13	48
20	T	87/258 (34%)	81 (93%)	6 (7%)	19	56
21	U	149/208 (72%)	142 (95%)	7 (5%)	32	68
22	V	84/144 (58%)	80 (95%)	4 (5%)	31	67
23	W	113/132 (86%)	108 (96%)	5 (4%)	35	69
24	X	87/117 (74%)	84 (97%)	3 (3%)	44	75
25	Y	512/707 (72%)	504 (98%)	8 (2%)	70	88
26	Z	412/737 (56%)	375 (91%)	37 (9%)	12	44
27	a	57/468 (12%)	57 (100%)	0	100	100
28	b	57/66 (86%)	57 (100%)	0	100	100
29	c	136/299 (46%)	118 (87%)	18 (13%)	5	28
30	d	107/260 (41%)	104 (97%)	3 (3%)	51	78
31	e	91/108 (84%)	83 (91%)	8 (9%)	12	45
32	f	136/641 (21%)	135 (99%)	1 (1%)	88	94
33	g	162/363 (45%)	157 (97%)	5 (3%)	47	77
34	h	108/429 (25%)	91 (84%)	17 (16%)	3	21
35	i	113/295 (38%)	86 (76%)	27 (24%)	1	7
36	j	152/205 (74%)	143 (94%)	9 (6%)	24	61
37	k	25/25 (100%)	25 (100%)	0	100	100
All	All	6478/10395 (62%)	5848 (90%)	630 (10%)	15	40

5 of 630 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	1220	ARG
5	E	92	THR
34	h	132	THR
3	C	56	THR
4	D	13	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 114 such sidechains are listed below:

Mol	Chain	Res	Type
21	U	198	GLN
26	Z	320	ASN
34	h	133	GLN
23	W	74	ASN
25	Y	138	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no 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 [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.