wwPDB EM Map/Model Validation Report (i)

May 26, 2016 – 04:46 PM EDT

PDB ID	:	5FZ5
EMDB ID:	:	3383
Title	:	Transcription initiation complex structures elucidate DNA opening (CC)
Authors	:	Plaschka, C.; Hantsche, M.; Dienemann, C.; Burzinski, C.; Plitzko, J.; Cramer, P.
Deposited on	:	2016-03-10
Resolution	:	8.80 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

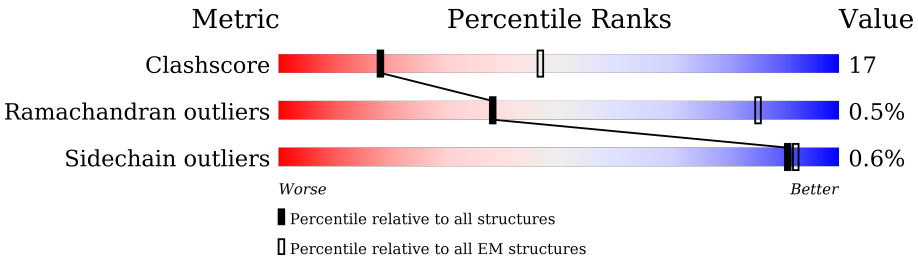
MolProbity	:	4.02b-467
Mogul	:	unknown
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20027674

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
ELECTRON MICROSCOPY

The reported resolution of this entry is 8.80 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.



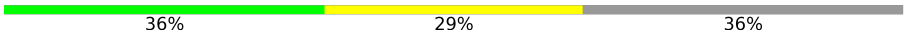
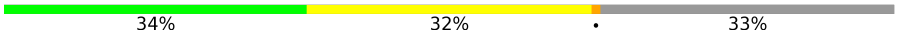
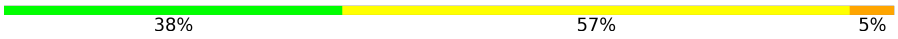
Metric	Whole archive (#Entries)	EM structures (#Entries)
Clashscore	114402	924
Ramachandran outliers	111179	726
Sidechain outliers	111093	686

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	1733	57% 23% 19%
2	B	1224	62% 32% 5%
3	C	318	53% 30% 18%
4	D	221	46% 25% 29%
5	E	215	76% 23% .
6	F	155	42% 12% 46%
7	G	171	51% 48% .
8	H	146	51% 40% .. 7%
9	I	122	61% 34% 5%

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	345	
14	N	56	
15	O	240	
16	Q	735	
17	R	400	
18	T	85	
19	U	286	
20	V	122	
21	W	482	
22	X	328	

2 Entry composition

There are 24 unique types of molecules in this entry. The entry contains 42144 atoms, of which 0 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
			Total	C	N	O	S		
1	A	1398	10997	6931	1927	2078	61	0	0

- Molecule 2 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
2	B	1157	9200	5819	1613	1713	55	0	0

- Molecule 3 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
3	C	262	2061	1299	343	406	13	0	0

- Molecule 4 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
4	D	157	1253	779	220	252	2	0	0

- Molecule 5 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1.

Mol	Chain	Residues	Atoms					AltConf	Trace
			Total	C	N	O	S		
5	E	213	1744	1107	308	318	11	0	0

- Molecule 6 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2.

Mol	Chain	Residues	Atoms					AltConf	Trace
6	F	83	Total	C	N	O	S	0	0
			670	428	114	125	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 RPABC 3.

Mol	Chain	Residues	Atoms					AltConf	Trace
8	H	136	Total	C	N	O	S	0	0
			1089	686	184	215	4		

- Molecule 9 is a protein called DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9.

Mol	Chain	Residues	Atoms					AltConf	Trace
9	I	116	Total	C	N	O	S	0	0
			944	581	172	181	10		

- Molecule 10 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5.

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	112	Total	C	N	O	S	0	0
			904	580	154	168	2		

- Molecule 12 is a protein called DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4.

Mol	Chain	Residues	Atoms					AltConf	Trace
12	L	45	Total	C	N	O	S	0	0
			358	221	71	62	4		

- Molecule 13 is a protein called TRANSCRIPTION INITIATION FACTOR IIB.

Mol	Chain	Residues	Atoms					AltConf	Trace
13	M	231	Total	C	N	O	S	0	0
			1785	1145	299	326	15		

- Molecule 14 is a DNA chain called SYNTHETIC CLOSED PROMOTER DNA CONSTRUCT.

Mol	Chain	Residues	Atoms					AltConf	Trace
14	N	56	Total	C	N	O	P	0	0
			1134	548	208	322	56		

- Molecule 15 is a protein called TATA-BOX-BINDING PROTEIN.

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

- Molecule 16 is a protein called TRANSCRIPTION INITIATION FACTOR IIF SUBUNIT ALPHA.

Mol	Chain	Residues	Atoms					AltConf	Trace
16	Q	148	Total	C	N	O	S	0	0
			1141	731	195	212	3		

- Molecule 17 is a protein called TRANSCRIPTION INITIATION FACTOR IIF SUBUNIT BETA.

Mol	Chain	Residues	Atoms					AltConf	Trace
17	R	199	Total	C	N	O	S	0	0
			1347	838	247	255	7		

- Molecule 18 is a DNA chain called SYNTHETIC CLOSED PROMOTER DNA CONSTRUCT.

Mol	Chain	Residues	Atoms					AltConf	Trace
18	T	56	Total	C	N	O	P	0	0
			1134	548	208	322	56		

- Molecule 19 is a protein called TRANSCRIPTION INITIATION FACTOR IIA LARGE SUBUNIT.

Mol	Chain	Residues	Atoms					AltConf	Trace
19	U	92	Total	C	N	O	S	0	0
			757	474	130	150	3		

- Molecule 20 is a protein called TRANSCRIPTION INITIATION FACTOR IIA SUBUNIT

2.

Mol	Chain	Residues	Atoms					AltConf	Trace
20	V	100	Total	C	N	O	S	0	0
			782	492	130	156	4		

- Molecule 21 is a protein called TRANSCRIPTION INITIATION FACTOR IIE SUBUNIT ALPHA.

Mol	Chain	Residues	Atoms				AltConf	Trace
21	W	168	Total	C	N	O	0	0
			835	499	168	168		

- Molecule 22 is a protein called TRANSCRIPTION INITIATION FACTOR IIE SUBUNIT BETA.

Mol	Chain	Residues	Atoms				AltConf	Trace
22	X	143	Total	C	N	O	0	0
			710	424	143	143		

- Molecule 23 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		AltConf
23	J	1	Total	Zn	0
			1	1	
23	B	1	Total	Zn	0
			1	1	
23	I	2	Total	Zn	0
			2	2	
23	C	1	Total	Zn	0
			1	1	
23	W	1	Total	Zn	0
			1	1	
23	A	2	Total	Zn	0
			2	2	
23	L	1	Total	Zn	0
			1	1	
23	M	1	Total	Zn	0
			1	1	

- Molecule 24 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

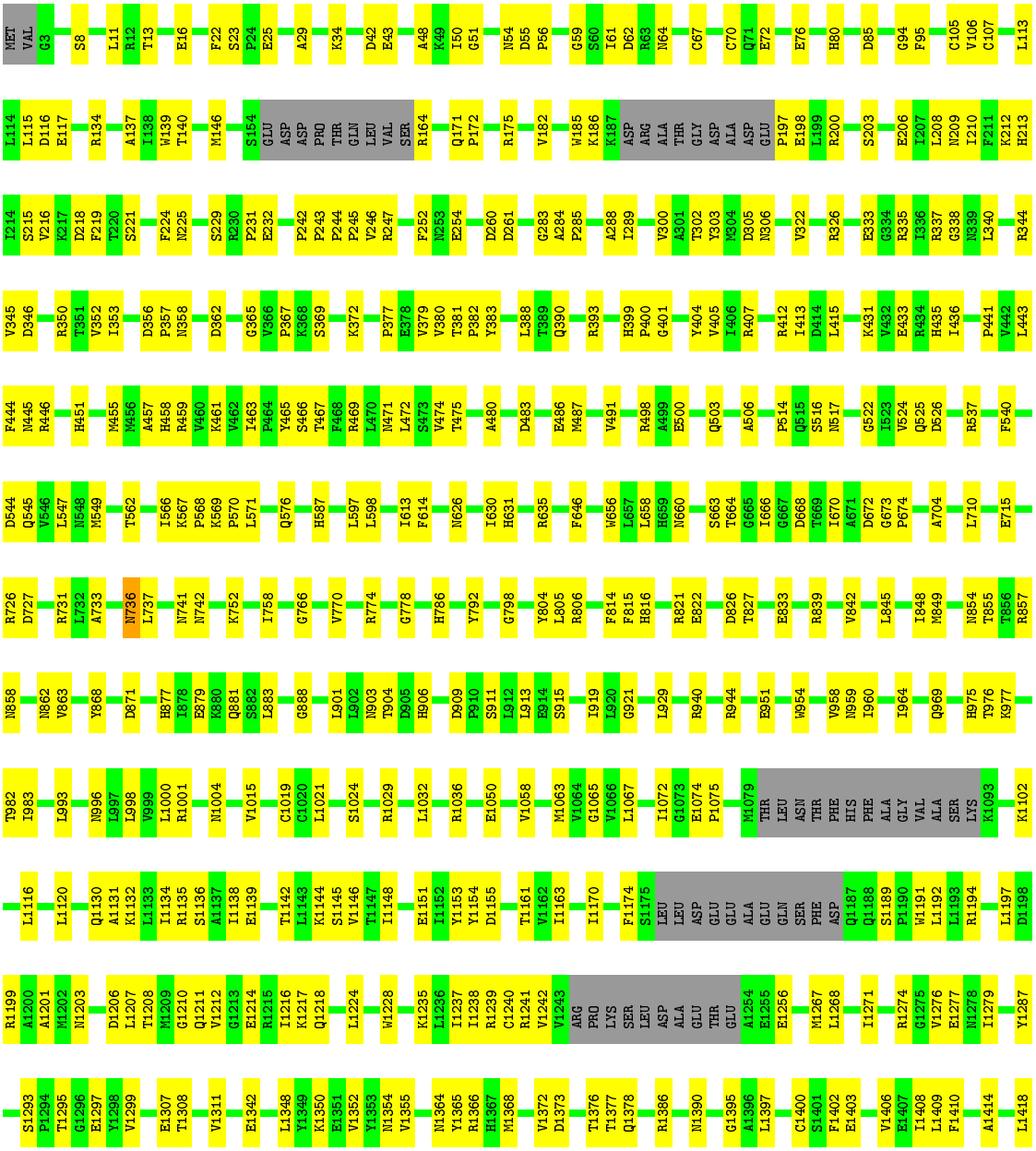
Mol	Chain	Residues	Atoms		AltConf
24	A	1	Total	Mg	0
			1	1	

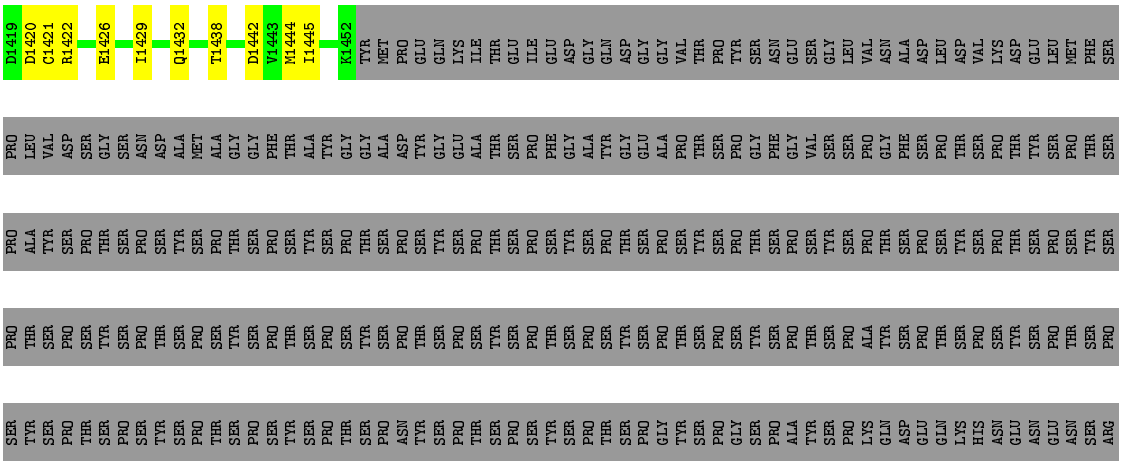
3 Residue-property plots (i)

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

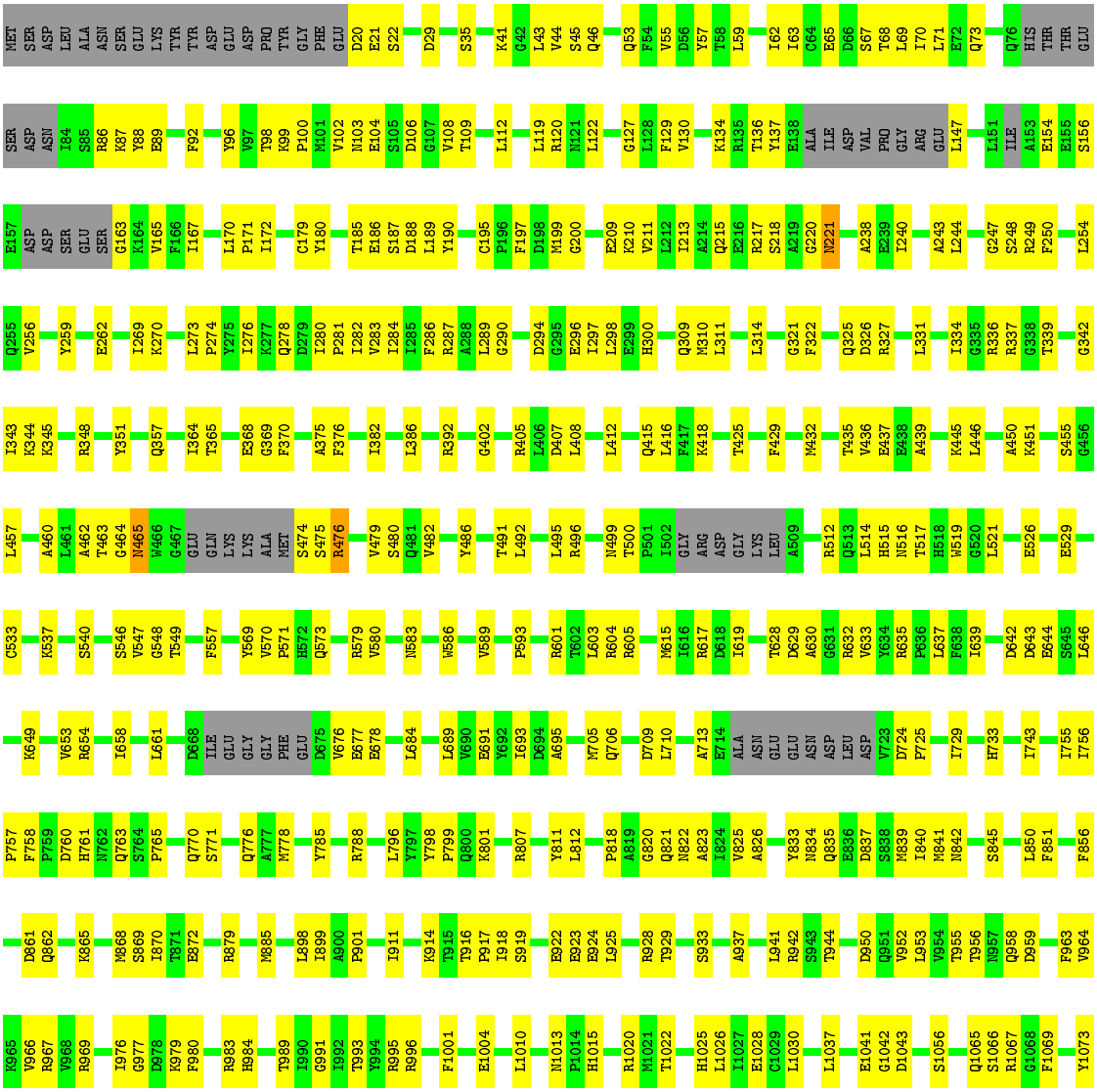
Chain A: 57% 23% 19%

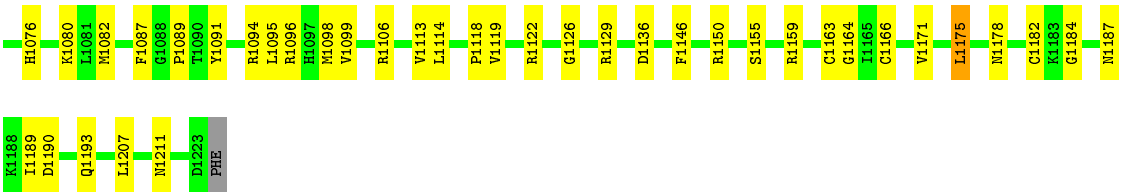




• Molecule 2: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB2

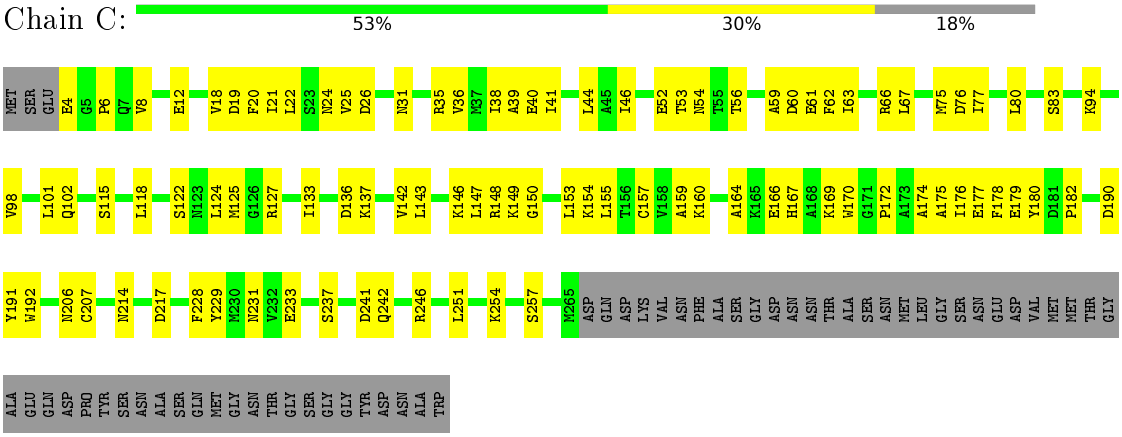
Chain B:





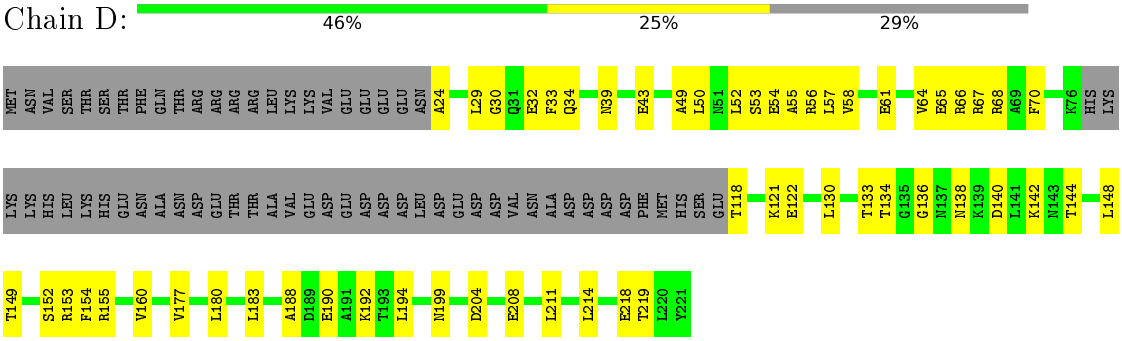
• Molecule 3: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB3

Chain C:



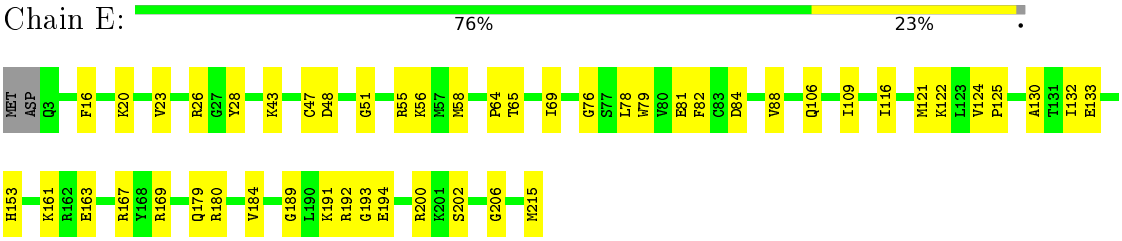
• Molecule 4: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB4

Chain D:



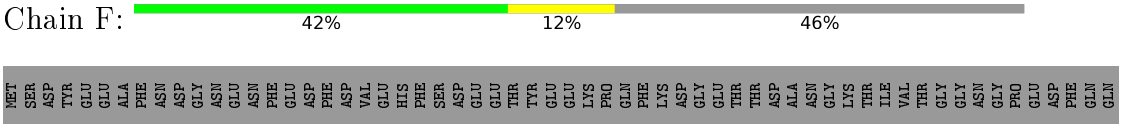
• Molecule 5: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 1

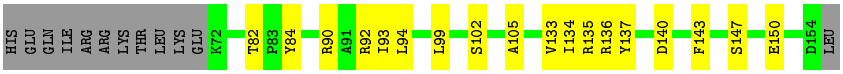
Chain E:



• Molecule 6: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 2

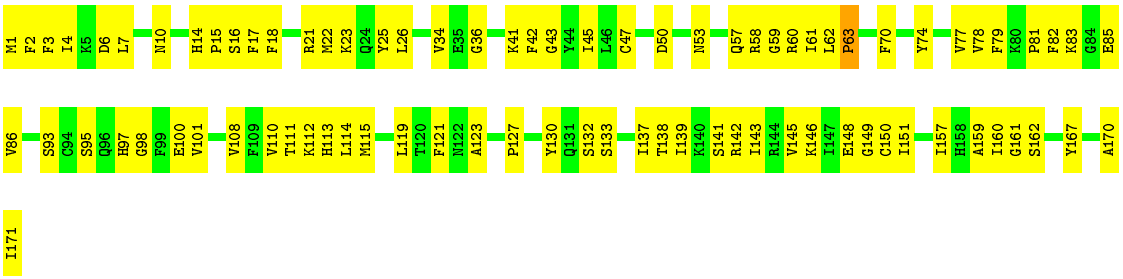
Chain F:





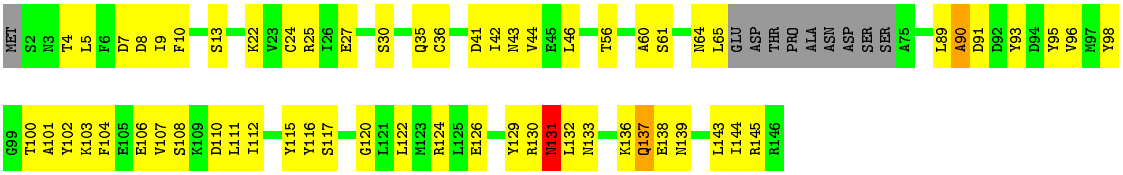
• Molecule 7: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB7

Chain G:



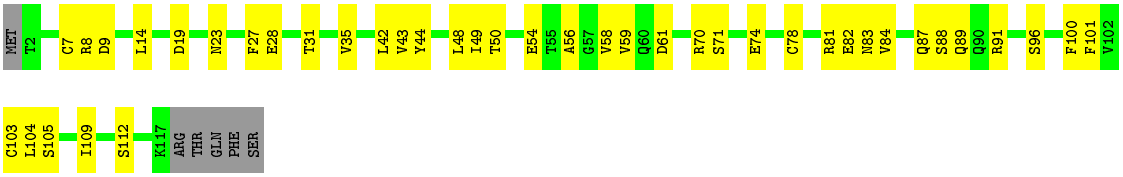
• Molecule 8: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 3

Chain H:



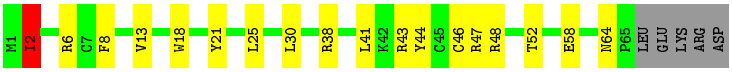
• Molecule 9: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB9

Chain I:



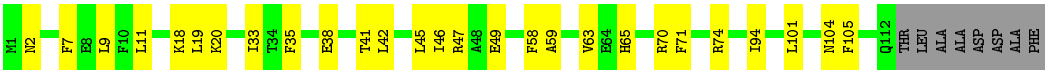
• Molecule 10: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 5

Chain J:



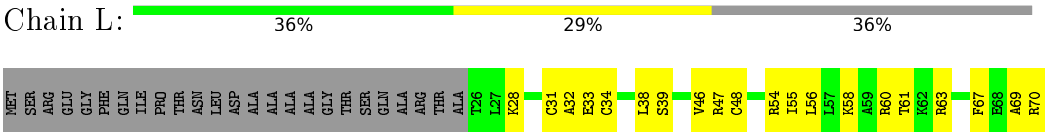
• Molecule 11: DNA-DIRECTED RNA POLYMERASE II SUBUNIT RPB11

Chain K:



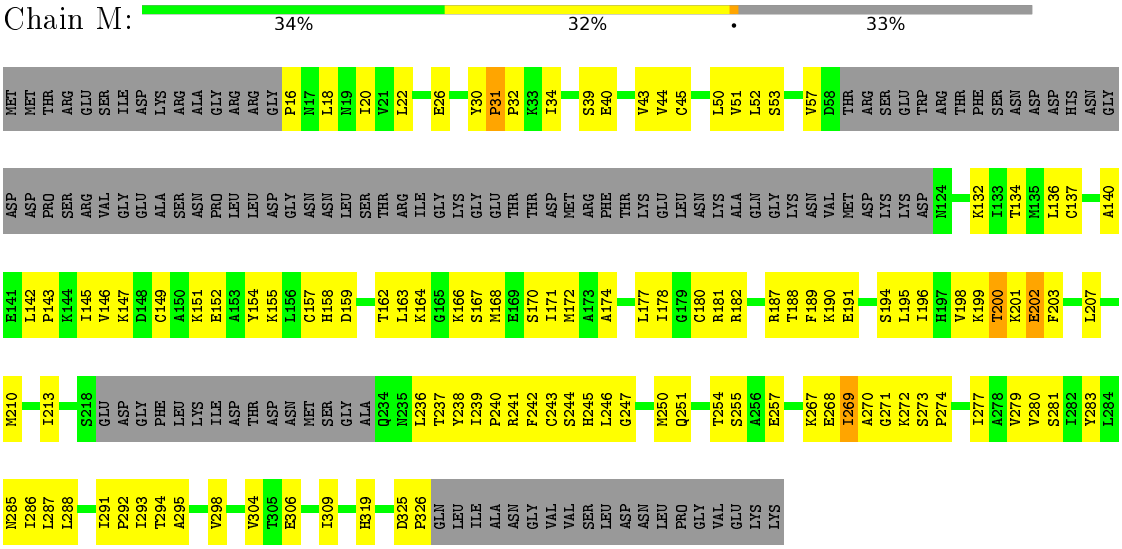
• Molecule 12: DNA-DIRECTED RNA POLYMERASES I, II, AND III SUBUNIT RPABC 4

Chain L:



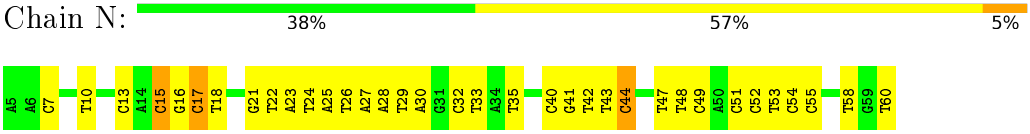
• Molecule 13: TRANSCRIPTION INITIATION FACTOR IIB

Chain M:



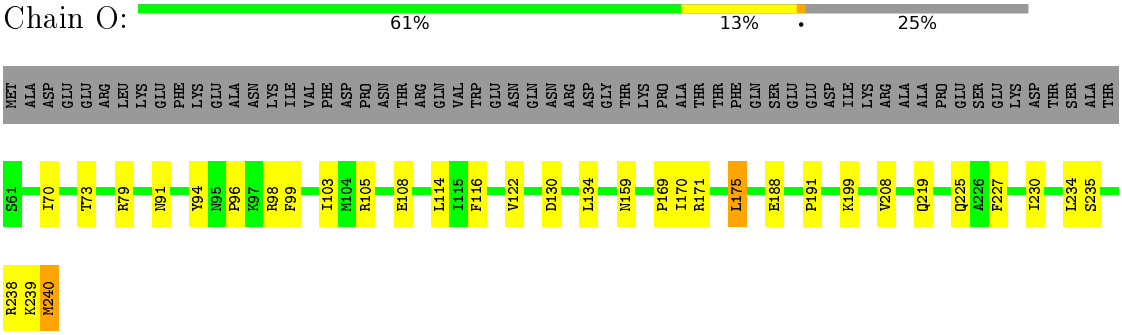
• Molecule 14: SYNTHETIC CLOSED PROMOTER DNA CONSTRUCT

Chain N:



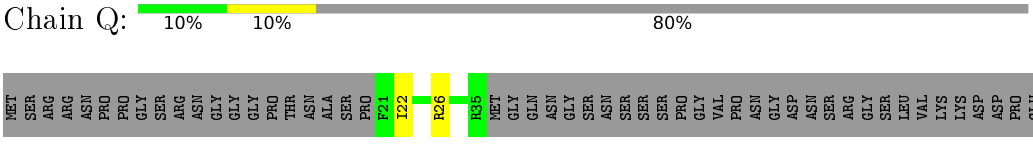
• Molecule 15: TATA-BOX-BINDING PROTEIN

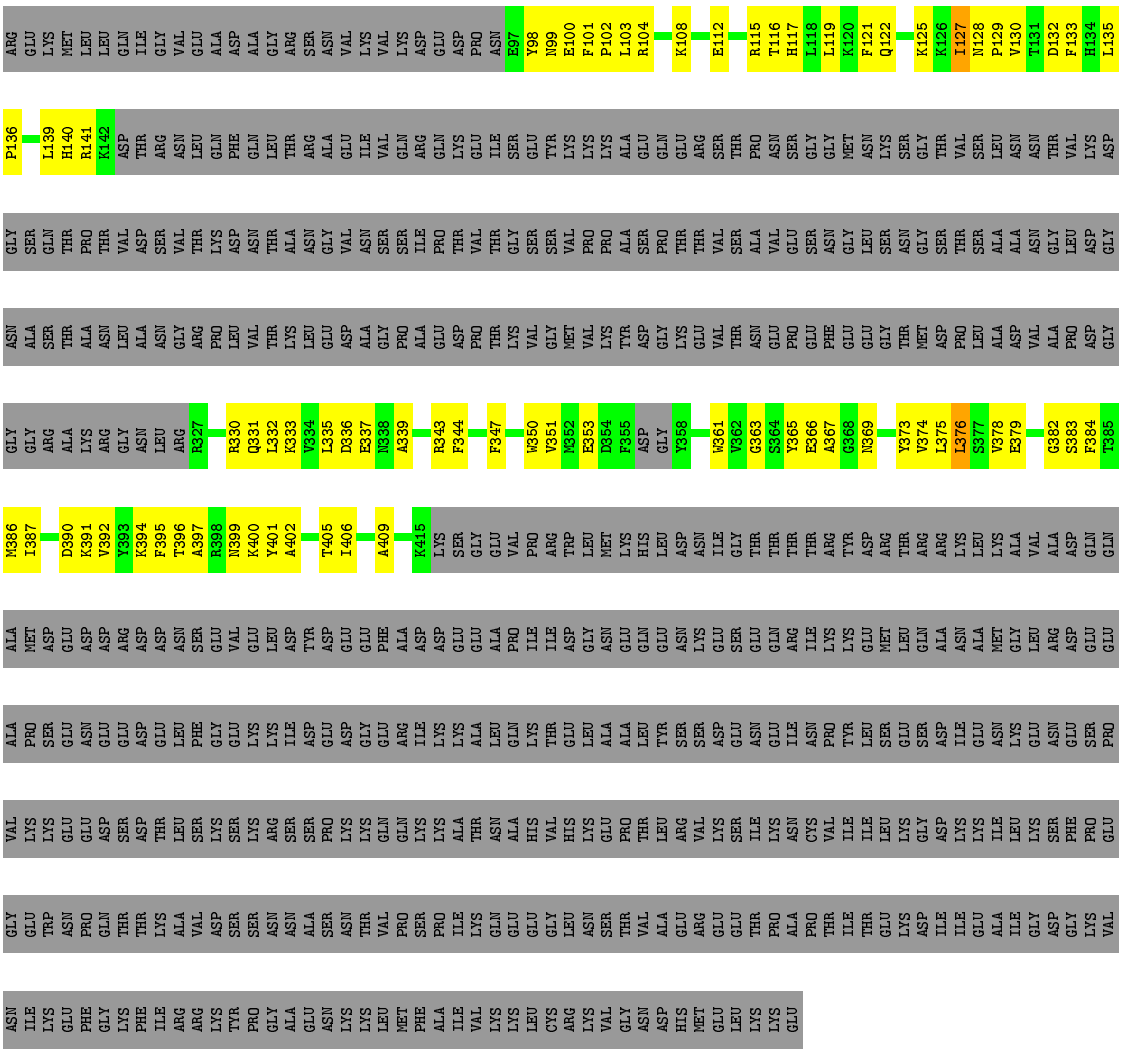
Chain O:



• Molecule 16: TRANSCRIPTION INITIATION FACTOR IIF SUBUNIT ALPHA

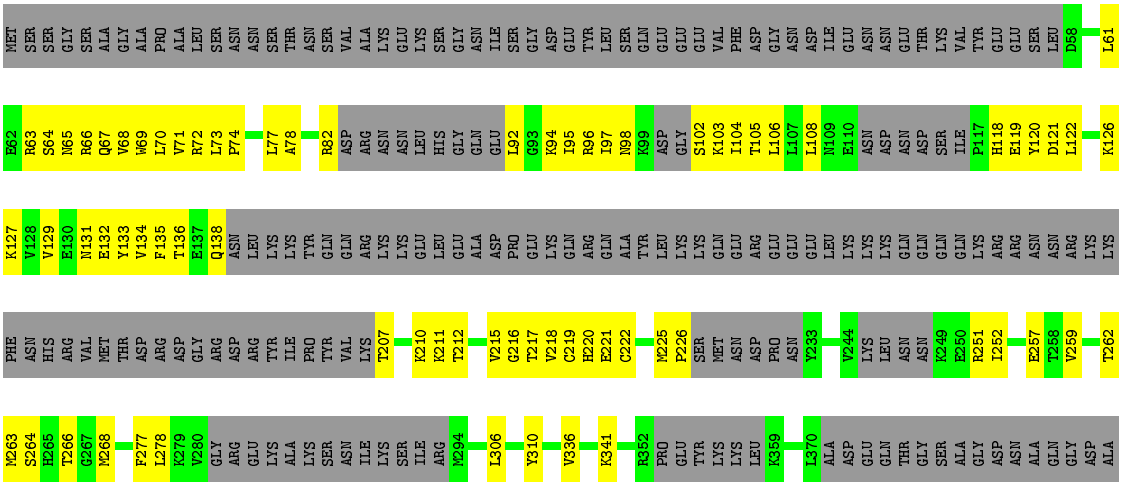
Chain Q:





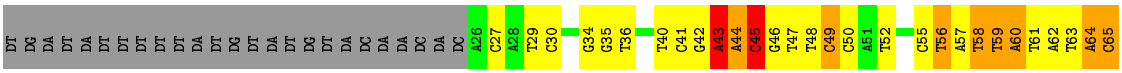
• Molecule 17: TRANSCRIPTION INITIATION FACTOR IIF SUBUNIT BETA

Chain R:



• Molecule 18: SYNTHETIC CLOSED PROMOTER DNA CONSTRUCT

Chain T:

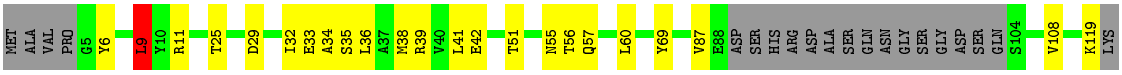


• Molecule 19: TRANSCRIPTION INITIATION FACTOR IIA LARGE SUBUNIT

Chain U:

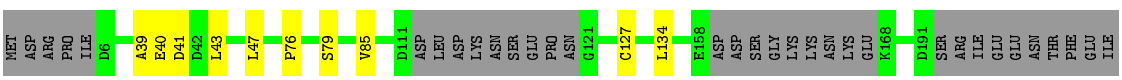
• Molecule 20: TRANSCRIPTION INITIATION FACTOR IIA SUBUNIT 2

Chain V:



• Molecule 21: TRANSCRIPTION INITIATION FACTOR IIE SUBUNIT ALPHA

Chain W:



MET ASP SER ILE ASN PRO ASP ASN GLU PRO ALA GLN GLU THR SER TYR GLN ASN ASN ARG THR LEU THR GLU GLN GLU MET GLU GLU ARG GLU ASN GLU LYS THR LEU ASN ASP TYR TYR ALA ALA LEU ALA LYS LYS GLN ALA LYS LEU ASN LYS GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU GLU MET GLU ASP VAL MET ASP ASP ASN ASP GLU THR ALA ARG GLU ASN ALA LEU GLU ASP GLU PHE GLU ASP VAL THR ASP THR ALA GLY THR ALA LYS THR GLU SER ASN THR SER ASN ASP VAL LYS GLN GLU SER ILE ASN ASP LYS THR GLU ASP ALA VAL ASN ALA THR ALA THR ALA SER GLY PRO SER ALA ASN ALA LYS PRO ASN ASP GLY ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP ASP GLU MET ASP ILE GLU PHE GLU ASP VAL

- Molecule 22: TRANSCRIPTION INITIATION FACTOR IIE SUBUNIT BETA

Chain X: 41% 56%

MET SER LYS ASN ARG ASP PRO LEU LEU ALA ASN LEU ASN ALA PHE LYS SER LYS VAL LYS SER ALA PRO VAL ILE ALA PRO ALA LYS VAL GLY GLN LYS LYS THR ASN ASP THR VAL ILE THR ILE ASP GLY ASN THR ARG LYS ARG THR ALA SER GLU ARG ALA GLN GLU ASN THR LEU ASN SER ALA LYS ASN PRO VAL LEU VAL ASP ILE LYS LYS GLU ALA GLY SER ASN SER SER ASN ALA ILE SER LEU ASP ASP ASP ASP ASP GLU ASP PHE GLY SER SER PRO SER LYS LYS VAL ARG PRO GLY SER ILE ALA ALA ALA LEU GLN ALA ASN GLN THR ASP ILE SER LYS SER H123 R168 L181 SER THR TYR ASP VAL H187 E225 L231 V232 G249 GLY ASN LEU LYS CYS ILE D256 K261 N265 V266 Q267 LEU PRO GLN PHE A272 D280 LEU GLY LEU LYS PRO ALA SER VAL ASP PRO ALA THR ILE LYS ARG GLN THR LYS ARG VAL GLU VAL LYS LYS LYS ARG GLN ARG LYS GLY LYS ILE THR ASN THR HIS MET THR GLY ILE LEU LYS ASP TYR SER HIS ARG VAL

4 Experimental information (i)

Property	Value	Source
Reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	Depositor
Number of particles used	Not provided	Depositor
Resolution determination method	Not provided	Depositor
CTF correction method	EACH PARTICLE	Depositor
Microscope	OTHER	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/Å^2$)	40	Depositor
Minimum defocus (nm)	800	Depositor
Maximum defocus (nm)	5000	Depositor
Magnification	37000	Depositor
Image detector	GATAN K2 SUMMIT (4K X 4K)	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	0.46	0/11192	0.55	1/15128 (0.0%)
10	J	0.53	0/541	0.57	0/727
11	K	0.43	0/922	0.53	0/1244
12	L	0.40	0/360	0.60	0/478
13	M	0.40	0/1809	0.60	0/2435
14	N	1.00	20/1273 (1.6%)	1.23	2/1942 (0.1%)
15	O	0.59	0/1443	0.78	1/1942 (0.1%)
16	Q	0.53	0/1165	0.68	1/1576 (0.1%)
17	R	0.43	0/1354	0.68	1/1832 (0.1%)
18	T	1.21	26/1273 (2.0%)	1.46	18/1942 (0.9%)
19	U	0.40	0/766	0.61	0/1032
2	B	0.51	0/9378	0.59	1/12646 (0.0%)
20	V	0.38	0/789	0.62	1/1066 (0.1%)
21	W	0.35	1/832 (0.1%)	0.47	0/1157
22	X	0.26	0/706	0.47	0/979
3	C	0.52	0/2099	0.57	0/2845
4	D	0.25	0/1262	0.44	0/1693
5	E	0.40	0/1780	0.49	0/2395
6	F	0.47	0/682	0.52	0/922
7	G	0.31	0/1368	0.50	0/1844
8	H	0.60	1/1107 (0.1%)	0.97	6/1499 (0.4%)
9	I	0.40	0/962	0.50	0/1295
All	All	0.53	48/43063 (0.1%)	0.67	32/58619 (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
10	J	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
13	M	0	1
18	T	0	2
2	B	0	1
8	H	0	2
All	All	0	7

The worst 5 of 48 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	H	131	ASN	C-N	-13.03	1.04	1.34
18	T	66	DA	P-O5'	-11.67	1.48	1.59
18	T	60	DA	C1'-N9	-10.36	1.32	1.47
18	T	44	DA	O3'-P	7.31	1.70	1.61
18	T	69	DG	C1'-N9	-6.83	1.37	1.47

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
8	H	131	ASN	N-CA-C	16.18	154.70	111.00
8	H	131	ASN	O-C-N	-15.87	97.31	122.70
8	H	131	ASN	C-N-CA	15.69	160.92	121.70
18	T	59	DT	O4'-C4'-C3'	-11.78	98.93	106.00
17	R	336	VAL	CB-CA-C	-10.95	90.59	111.40

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	1175	LEU	Peptide
8	H	131	ASN	Mainchain,Peptide
10	J	2	ILE	Peptide
13	M	292	PRO	Peptide
18	T	43	DA	Sidechain

5.2 Too-close contacts (i)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	10997	0	11082	294	0
2	B	9200	0	9194	376	0
3	C	2061	0	2029	70	0
4	D	1253	0	1273	61	0
5	E	1744	0	1772	31	0
6	F	670	0	690	15	0
7	G	1340	0	1356	96	0
8	H	1089	0	1061	40	0
9	I	944	0	899	30	0
10	J	532	0	543	19	0
11	K	904	0	911	21	0
12	L	358	0	380	26	0
13	M	1785	0	1891	165	0
14	N	1134	0	633	57	0
15	O	1416	0	1493	103	0
16	Q	1141	0	1027	99	0
17	R	1347	0	1130	89	0
18	T	1134	0	633	124	0
19	U	757	0	747	21	0
20	V	782	0	790	17	0
21	W	835	0	348	12	0
22	X	710	0	287	5	0
23	A	2	0	0	0	0
23	B	1	0	0	0	0
23	C	1	0	0	0	0
23	I	2	0	0	0	0
23	J	1	0	0	0	0
23	L	1	0	0	0	0
23	M	1	0	0	0	0
23	W	1	0	0	0	0
24	A	1	0	0	0	0
All	All	42144	0	40169	1424	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 17.

The worst 5 of 1424 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:405:ARG:NH1	2:B:632:ARG:HG2	1.23	1.42
2:B:865:LYS:CE	13:M:145:ILE:HD11	1.46	1.42
1:A:197:PRO:HD3	18:T:35:DG:OP2	1.22	1.36

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:868:MET:HB2	13:M:149:CYS:SG	1.64	1.35
13:M:267:LYS:HE3	15:O:208:VAL:CG1	1.59	1.31

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 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	1386/1733 (80%)	1306 (94%)	77 (6%)	3 (0%)	52 86
2	B	1139/1224 (93%)	1085 (95%)	49 (4%)	5 (0%)	39 80
3	C	260/318 (82%)	246 (95%)	14 (5%)	0	100 100
4	D	153/221 (69%)	145 (95%)	8 (5%)	0	100 100
5	E	211/215 (98%)	209 (99%)	2 (1%)	0	100 100
6	F	81/155 (52%)	76 (94%)	5 (6%)	0	100 100
7	G	169/171 (99%)	160 (95%)	8 (5%)	1 (1%)	30 74
8	H	132/146 (90%)	114 (86%)	14 (11%)	4 (3%)	5 42
9	I	114/122 (93%)	105 (92%)	9 (8%)	0	100 100
10	J	63/70 (90%)	57 (90%)	5 (8%)	1 (2%)	12 56
11	K	110/120 (92%)	107 (97%)	3 (3%)	0	100 100
12	L	43/70 (61%)	37 (86%)	6 (14%)	0	100 100
13	M	225/345 (65%)	202 (90%)	20 (9%)	3 (1%)	15 60
15	O	178/240 (74%)	174 (98%)	4 (2%)	0	100 100
16	Q	140/735 (19%)	124 (89%)	13 (9%)	3 (2%)	9 50
17	R	181/400 (45%)	171 (94%)	9 (5%)	1 (1%)	30 74
19	U	88/286 (31%)	82 (93%)	3 (3%)	3 (3%)	5 40
20	V	96/122 (79%)	88 (92%)	7 (7%)	1 (1%)	19 65

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
21	W	162/482 (34%)	154 (95%)	7 (4%)	1 (1%)	30 74
22	X	135/328 (41%)	125 (93%)	9 (7%)	1 (1%)	26 71
All	All	5066/7503 (68%)	4767 (94%)	272 (5%)	27 (0%)	38 77

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	343	ILE
2	B	476	ARG
8	H	132	LEU
13	M	269	ILE
17	R	259	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 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	1221/1520 (80%)	1219 (100%)	2 (0%)	95 97
2	B	999/1061 (94%)	998 (100%)	1 (0%)	95 97
3	C	230/274 (84%)	230 (100%)	0	100 100
4	D	139/200 (70%)	139 (100%)	0	100 100
5	E	195/197 (99%)	195 (100%)	0	100 100
6	F	73/137 (53%)	73 (100%)	0	100 100
7	G	152/152 (100%)	152 (100%)	0	100 100
8	H	119/128 (93%)	119 (100%)	0	100 100
9	I	110/116 (95%)	110 (100%)	0	100 100
10	J	60/65 (92%)	60 (100%)	0	100 100
11	K	97/102 (95%)	97 (100%)	0	100 100
12	L	40/57 (70%)	40 (100%)	0	100 100
13	M	201/299 (67%)	200 (100%)	1 (0%)	92 96
15	O	152/205 (74%)	141 (93%)	11 (7%)	18 55

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
16	Q	108/641 (17%)	108 (100%)	0	100 100
17	R	107/363 (30%)	107 (100%)	0	100 100
19	U	84/260 (32%)	79 (94%)	5 (6%)	24 60
20	V	90/108 (83%)	85 (94%)	5 (6%)	26 62
All	All	4177/5885 (71%)	4152 (99%)	25 (1%)	91 95

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
15	O	219	GLN
15	O	235	SER
20	V	57	GLN
15	O	225	GLN
15	O	240	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 13 such sidechains are listed below:

Mol	Chain	Res	Type
13	M	197	HIS
15	O	95	ASN
19	U	280	GLN
4	D	173	HIS
16	Q	122	GLN

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 carbohydrates in this entry.

5.6 Ligand geometry

Of 11 ligands modelled in this entry, 11 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.