wwPDB X-ray Structure Validation Summary Report (i)

May 29, 2020 – 08:16 am BST

PDB ID	:	4YLP
Title	:	E. coli Transcription Initiation Complex - 16-bp spacer and 5-nt RNA
Authors	:	Zuo, Y.; Steitz, T.A.
Deposited on	:	2015-03-05
Resolution	:	5.50 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 following versions of software and data (see references (1)) were used in the production of this report:

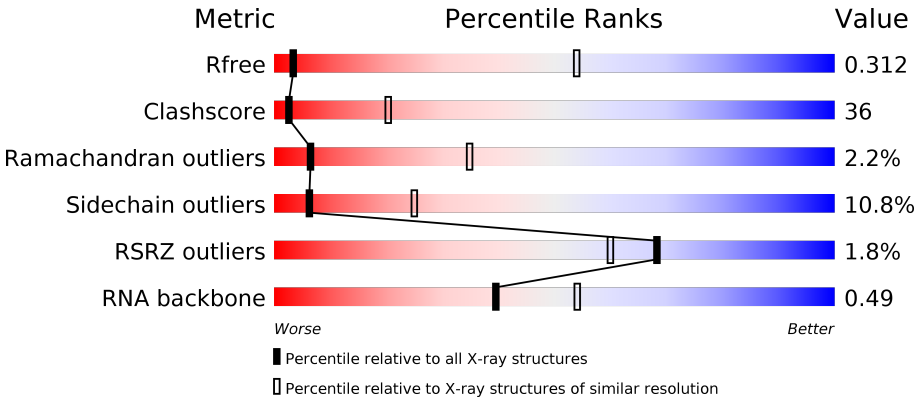
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 5.50 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	1019 (7.12-3.82)
Clashscore	141614	1010 (7.10-3.90)
Ramachandran outliers	138981	1014 (7.12-3.82)
Sidechain outliers	138945	1191 (7.20-3.80)
RSRZ outliers	127900	1023 (7.08-3.76)
RNA backbone	3102	1074 (7.80-3.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria 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	242	40% 44% 10% • 5%
1	B	242	39% 42% 13% 6%
1	G	242	49% 40% 5% • 5%
1	H	242	% 51% 36% 7% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	M	242	2% 52% 39% 5%
1	N	242	2% 52% 37% 5% 6%
2	C	1342	47% 48% 5%
2	I	1342	44% 49% 7%
2	O	1342	49% 45% 5%
3	D	1407	2% 45% 44% 7%
3	J	1407	2% 44% 44% 9%
3	P	1407	3% 45% 44% 7%
4	E	90	7% 59% 40%
4	K	90	9% 54% 39% 7%
4	Q	90	4% 57% 40%
5	F	628	4% 41% 32% 6% 21%
5	L	628	5% 45% 29% 21%
5	R	628	4% 42% 30% 7% 21%
6	1	49	31% 69%
6	4	49	4% 35% 63%
6	7	49	2% 37% 63%
7	2	49	39% 61%
7	5	49	35% 65%
7	8	49	41% 59%
8	3	5	80% 20%
8	6	5	40% 40% 20%
8	9	5	40% 60%

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
9	ZN	D	1502	-	-	X	-
9	ZN	P	1501	-	-	X	-
9	ZN	P	1502	-	-	X	-

2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 94668 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 alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	230	Total	C	N	O	S	0	0	0
			1787	1112	317	352	6			
1	B	228	Total	C	N	O	S	0	0	0
			1767	1100	312	349	6			
1	G	230	Total	C	N	O	S	0	0	0
			1787	1112	317	352	6			
1	H	228	Total	C	N	O	S	0	0	0
			1767	1100	312	349	6			
1	M	230	Total	C	N	O	S	0	0	0
			1787	1112	317	352	6			
1	N	228	Total	C	N	O	S	0	0	0
			1767	1100	312	349	6			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	ALA	-	expression tag	UNP A7ZSI4
A	-5	HIS	-	expression tag	UNP A7ZSI4
A	-4	HIS	-	expression tag	UNP A7ZSI4
A	-3	HIS	-	expression tag	UNP A7ZSI4
A	-2	HIS	-	expression tag	UNP A7ZSI4
A	-1	HIS	-	expression tag	UNP A7ZSI4
A	0	HIS	-	expression tag	UNP A7ZSI4
B	-6	ALA	-	expression tag	UNP A7ZSI4
B	-5	HIS	-	expression tag	UNP A7ZSI4
B	-4	HIS	-	expression tag	UNP A7ZSI4
B	-3	HIS	-	expression tag	UNP A7ZSI4
B	-2	HIS	-	expression tag	UNP A7ZSI4
B	-1	HIS	-	expression tag	UNP A7ZSI4
B	0	HIS	-	expression tag	UNP A7ZSI4
G	-6	ALA	-	expression tag	UNP A7ZSI4
G	-5	HIS	-	expression tag	UNP A7ZSI4
G	-4	HIS	-	expression tag	UNP A7ZSI4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	-3	HIS	-	expression tag	UNP A7ZSI4
G	-2	HIS	-	expression tag	UNP A7ZSI4
G	-1	HIS	-	expression tag	UNP A7ZSI4
G	0	HIS	-	expression tag	UNP A7ZSI4
H	-6	ALA	-	expression tag	UNP A7ZSI4
H	-5	HIS	-	expression tag	UNP A7ZSI4
H	-4	HIS	-	expression tag	UNP A7ZSI4
H	-3	HIS	-	expression tag	UNP A7ZSI4
H	-2	HIS	-	expression tag	UNP A7ZSI4
H	-1	HIS	-	expression tag	UNP A7ZSI4
H	0	HIS	-	expression tag	UNP A7ZSI4
M	-6	ALA	-	expression tag	UNP A7ZSI4
M	-5	HIS	-	expression tag	UNP A7ZSI4
M	-4	HIS	-	expression tag	UNP A7ZSI4
M	-3	HIS	-	expression tag	UNP A7ZSI4
M	-2	HIS	-	expression tag	UNP A7ZSI4
M	-1	HIS	-	expression tag	UNP A7ZSI4
M	0	HIS	-	expression tag	UNP A7ZSI4
N	-6	ALA	-	expression tag	UNP A7ZSI4
N	-5	HIS	-	expression tag	UNP A7ZSI4
N	-4	HIS	-	expression tag	UNP A7ZSI4
N	-3	HIS	-	expression tag	UNP A7ZSI4
N	-2	HIS	-	expression tag	UNP A7ZSI4
N	-1	HIS	-	expression tag	UNP A7ZSI4
N	0	HIS	-	expression tag	UNP A7ZSI4

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1341	Total	C	N	O	S	0	0	0
			10576	6636	1842	2055	43			
2	I	1341	Total	C	N	O	S	0	0	0
			10576	6636	1842	2055	43			
2	O	1341	Total	C	N	O	S	0	0	0
			10576	6636	1842	2055	43			

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1362	Total	C	N	O	S	0	0	0
			10568	6633	1887	1998	50			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	1362	Total	C	N	O	S	0	0	0
			10568	6633	1887	1998	50			
3	P	1362	Total	C	N	O	S	0	0	0
			10568	6633	1887	1998	50			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			
4	K	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			
4	Q	90	Total	C	N	O	S	0	0	0
			708	430	136	141	1			

- Molecule 5 is a protein called RNA polymerase sigma factor RpoD.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	497	Total	C	N	O	S	0	0	0
			4022	2512	719	768	23			
5	L	497	Total	C	N	O	S	0	0	0
			4022	2512	719	768	23			
5	R	497	Total	C	N	O	S	0	0	0
			4022	2512	719	768	23			

There are 45 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-14	MET	-	expression tag	UNP P00579
F	-13	ARG	-	expression tag	UNP P00579
F	-12	GLY	-	expression tag	UNP P00579
F	-11	SER	-	expression tag	UNP P00579
F	-10	HIS	-	expression tag	UNP P00579
F	-9	HIS	-	expression tag	UNP P00579
F	-8	HIS	-	expression tag	UNP P00579
F	-7	HIS	-	expression tag	UNP P00579
F	-6	HIS	-	expression tag	UNP P00579
F	-5	HIS	-	expression tag	UNP P00579
F	-4	THR	-	expression tag	UNP P00579
F	-3	ASP	-	expression tag	UNP P00579
F	-2	GLN	-	expression tag	UNP P00579
F	-1	PHE	-	expression tag	UNP P00579

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	0	THR	-	expression tag	UNP P00579
L	-14	MET	-	expression tag	UNP P00579
L	-13	ARG	-	expression tag	UNP P00579
L	-12	GLY	-	expression tag	UNP P00579
L	-11	SER	-	expression tag	UNP P00579
L	-10	HIS	-	expression tag	UNP P00579
L	-9	HIS	-	expression tag	UNP P00579
L	-8	HIS	-	expression tag	UNP P00579
L	-7	HIS	-	expression tag	UNP P00579
L	-6	HIS	-	expression tag	UNP P00579
L	-5	HIS	-	expression tag	UNP P00579
L	-4	THR	-	expression tag	UNP P00579
L	-3	ASP	-	expression tag	UNP P00579
L	-2	GLN	-	expression tag	UNP P00579
L	-1	PHE	-	expression tag	UNP P00579
L	0	THR	-	expression tag	UNP P00579
R	-14	MET	-	expression tag	UNP P00579
R	-13	ARG	-	expression tag	UNP P00579
R	-12	GLY	-	expression tag	UNP P00579
R	-11	SER	-	expression tag	UNP P00579
R	-10	HIS	-	expression tag	UNP P00579
R	-9	HIS	-	expression tag	UNP P00579
R	-8	HIS	-	expression tag	UNP P00579
R	-7	HIS	-	expression tag	UNP P00579
R	-6	HIS	-	expression tag	UNP P00579
R	-5	HIS	-	expression tag	UNP P00579
R	-4	THR	-	expression tag	UNP P00579
R	-3	ASP	-	expression tag	UNP P00579
R	-2	GLN	-	expression tag	UNP P00579
R	-1	PHE	-	expression tag	UNP P00579
R	0	THR	-	expression tag	UNP P00579

- Molecule 6 is a DNA chain called NT strand DNA (49-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	1	49	Total	C	N	O	P	0	0	0
			996	476	178	294	48			
6	4	49	Total	C	N	O	P	0	0	0
			996	476	178	294	48			
6	7	49	Total	C	N	O	P	0	0	0
			996	476	178	294	48			

- Molecule 7 is a DNA chain called T strand DNA (49-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	2	49	Total	C	N	O	P	0	0	0
			1012	481	191	292	48			
7	5	49	Total	C	N	O	P	0	0	0
			1012	481	191	292	48			
7	8	49	Total	C	N	O	P	0	0	0
			1012	481	191	292	48			

- Molecule 8 is a RNA chain called RNA (5'-R(*(GTP))-R(P*AP*GP*UP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	3	5	Total	C	N	O	P	0	0	0
			117	48	20	42	7			
8	6	5	Total	C	N	O	P	0	0	0
			117	48	20	42	7			
8	9	5	Total	C	N	O	P	0	0	0
			117	48	20	42	7			

- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	P	2	Total	Zn	0	0
			2	2		
9	J	2	Total	Zn	0	0
			2	2		
9	D	2	Total	Zn	0	0
			2	2		

- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

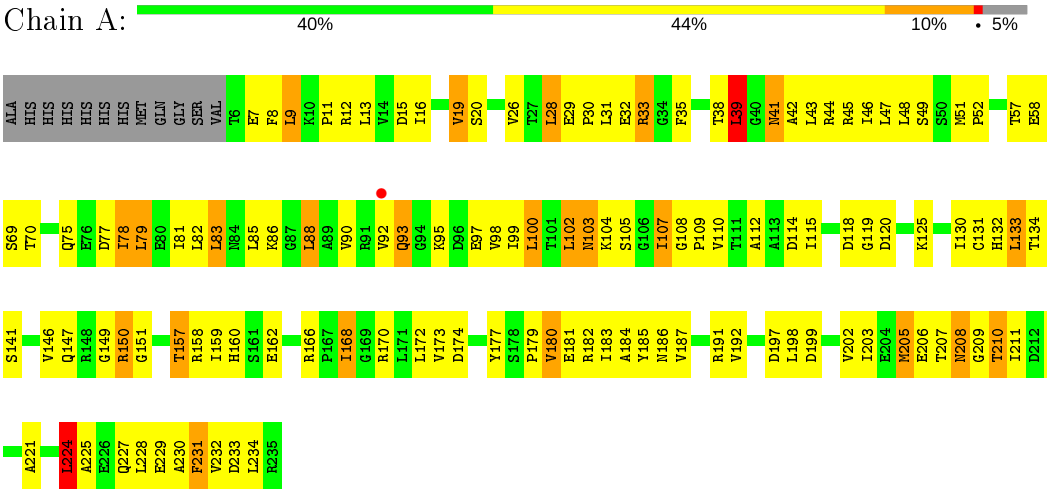
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	P	1	Total	Mg	0	0
			1	1		
10	J	1	Total	Mg	0	0
			1	1		
10	D	1	Total	Mg	0	0
			1	1		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

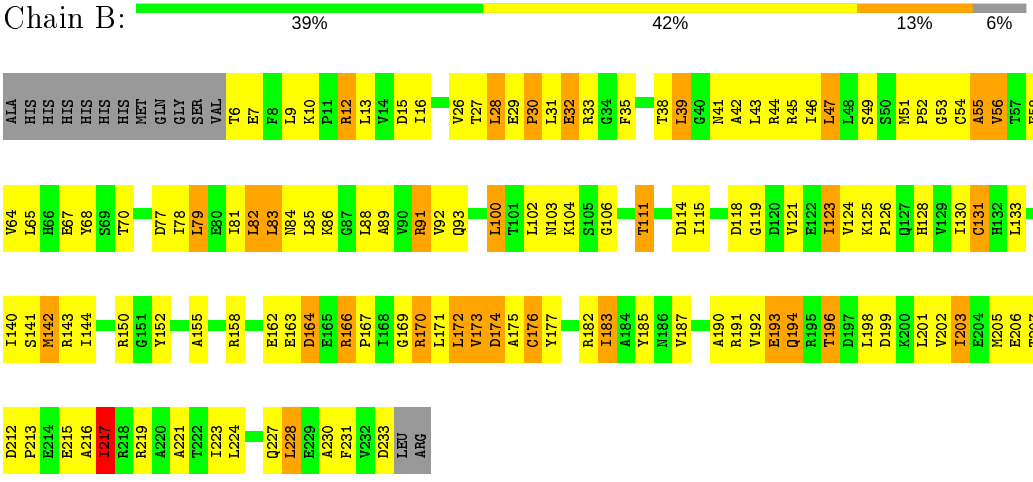
- Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A:



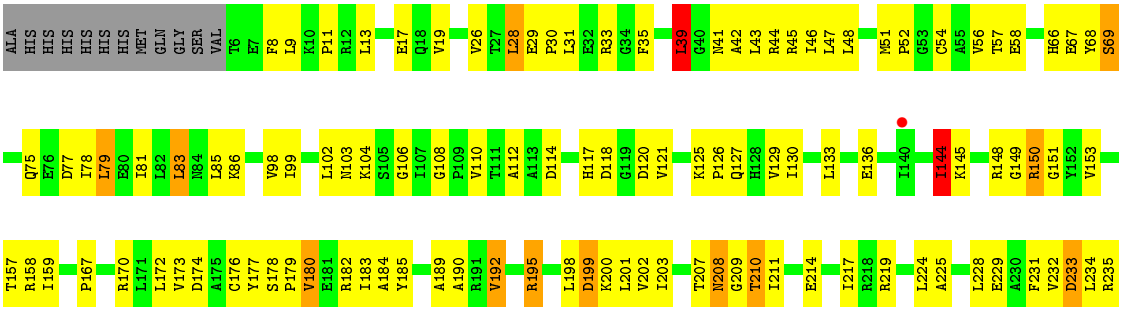
- Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain B:



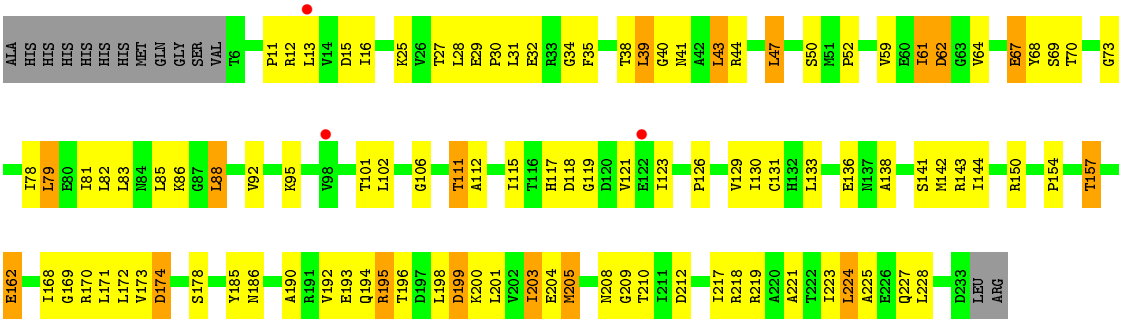
- Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain G:



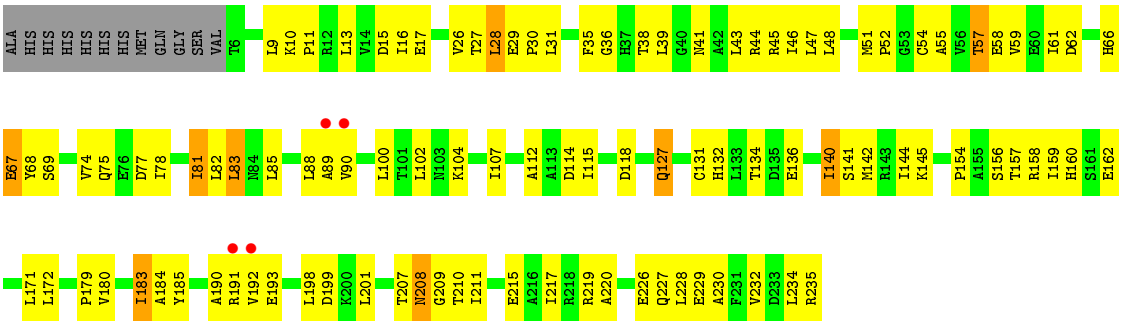
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain H:



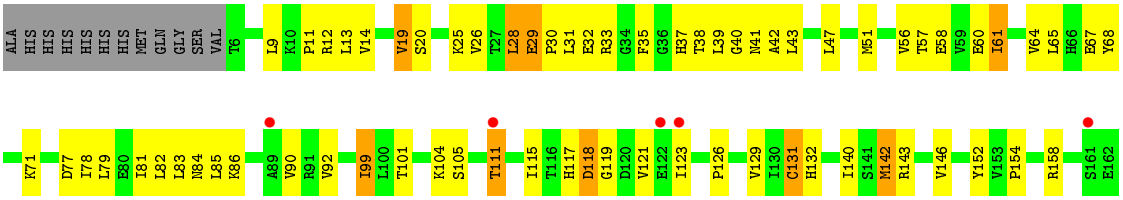
• Molecule 1: DNA-directed RNA polymerase subunit alpha

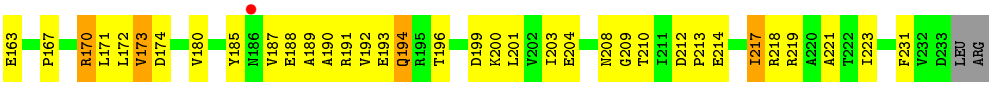
Chain M:



• Molecule 1: DNA-directed RNA polymerase subunit alpha

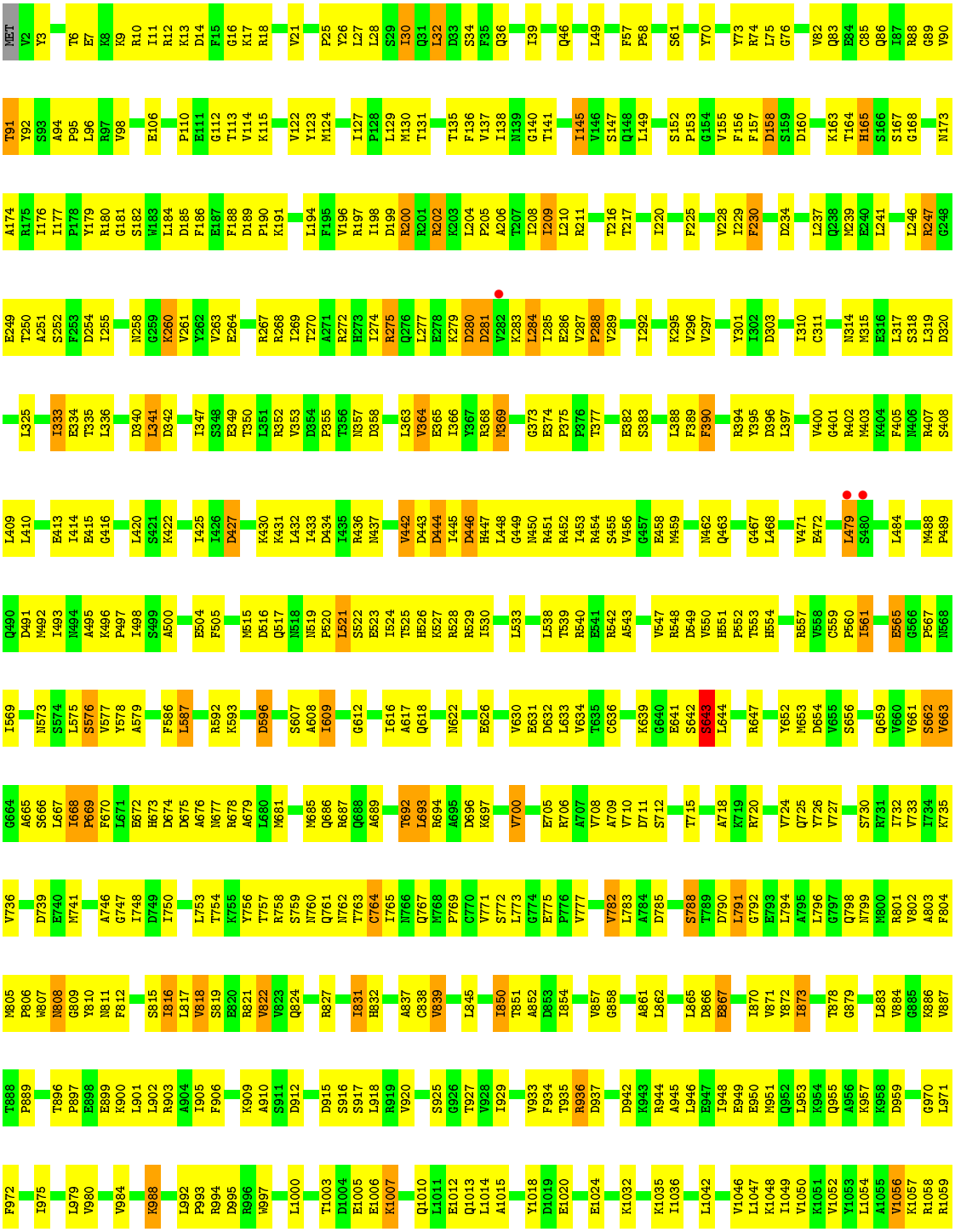
Chain N:

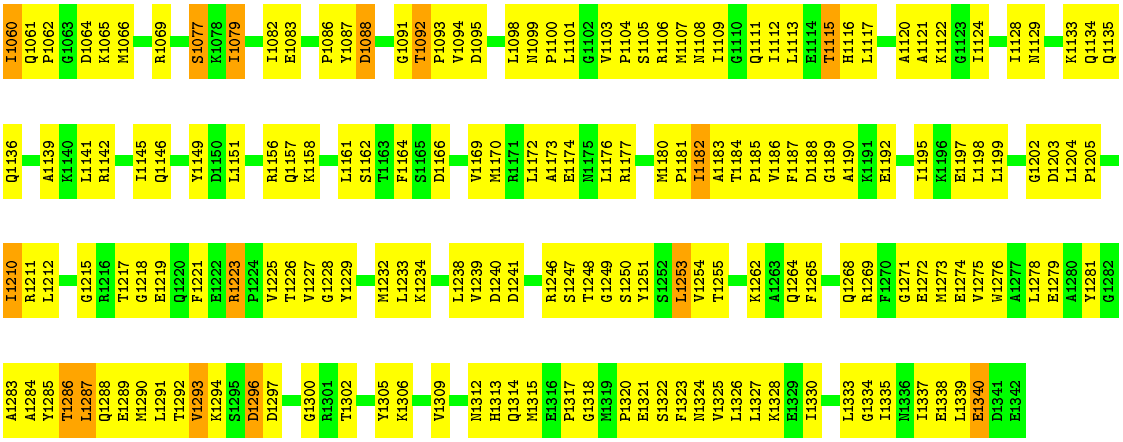




- Molecule 2: DNA-directed RNA polymerase subunit beta

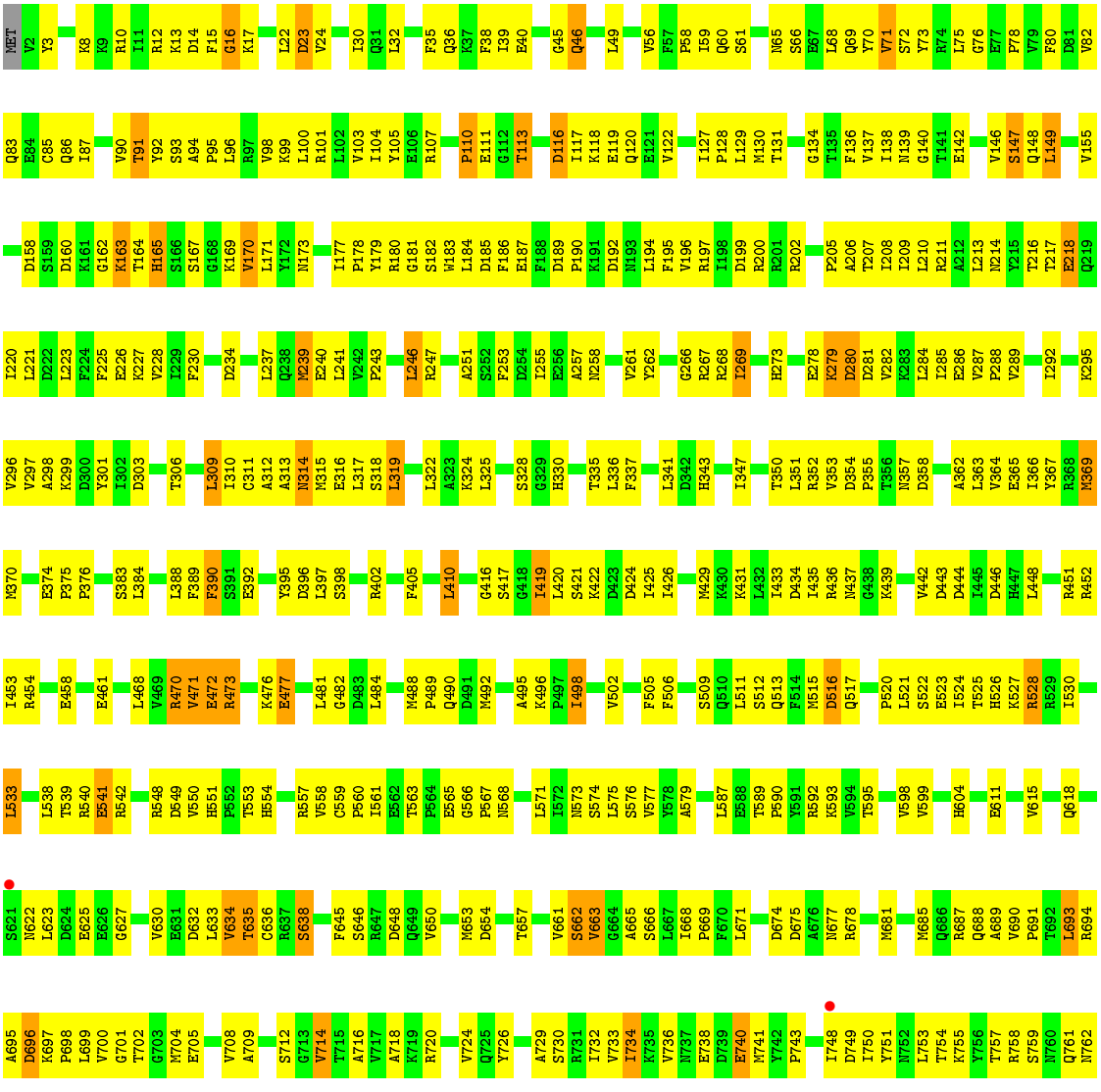
Chain C:

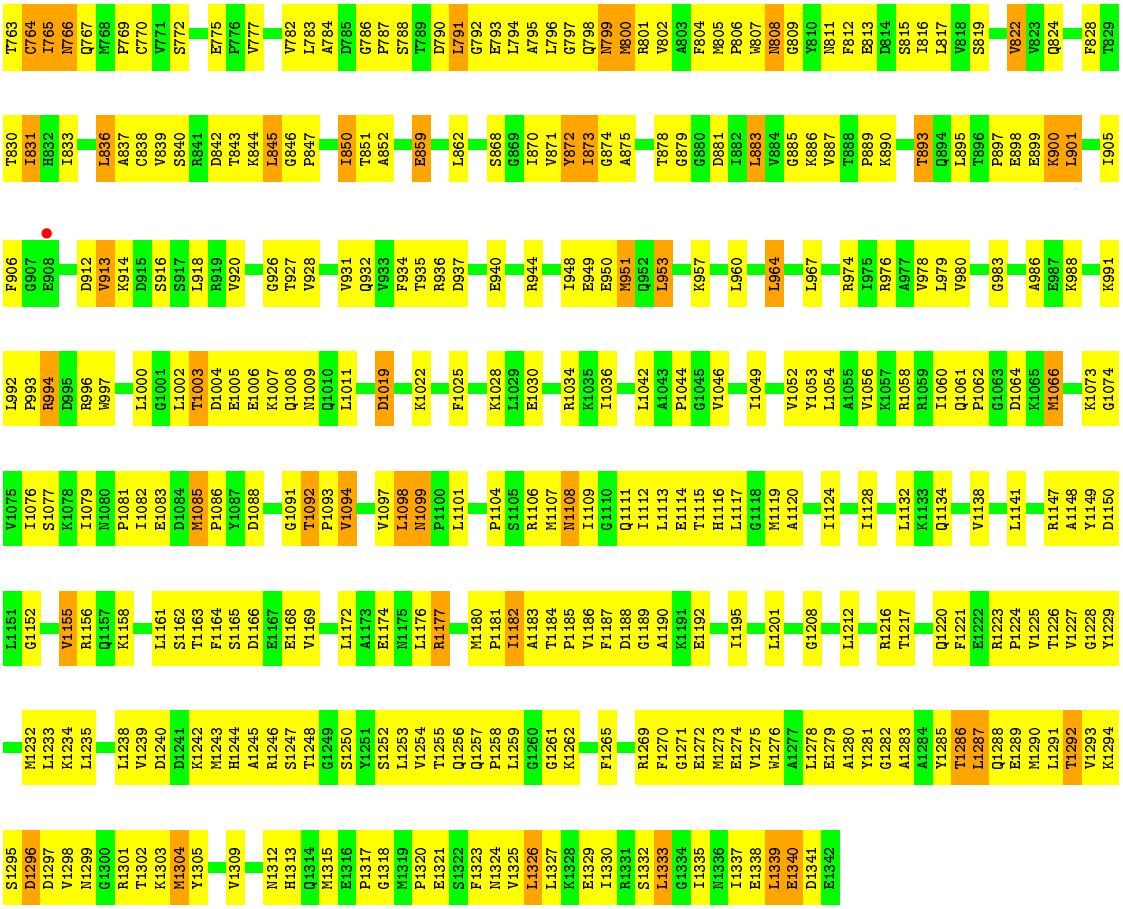




• Molecule 2: DNA-directed RNA polymerase subunit beta

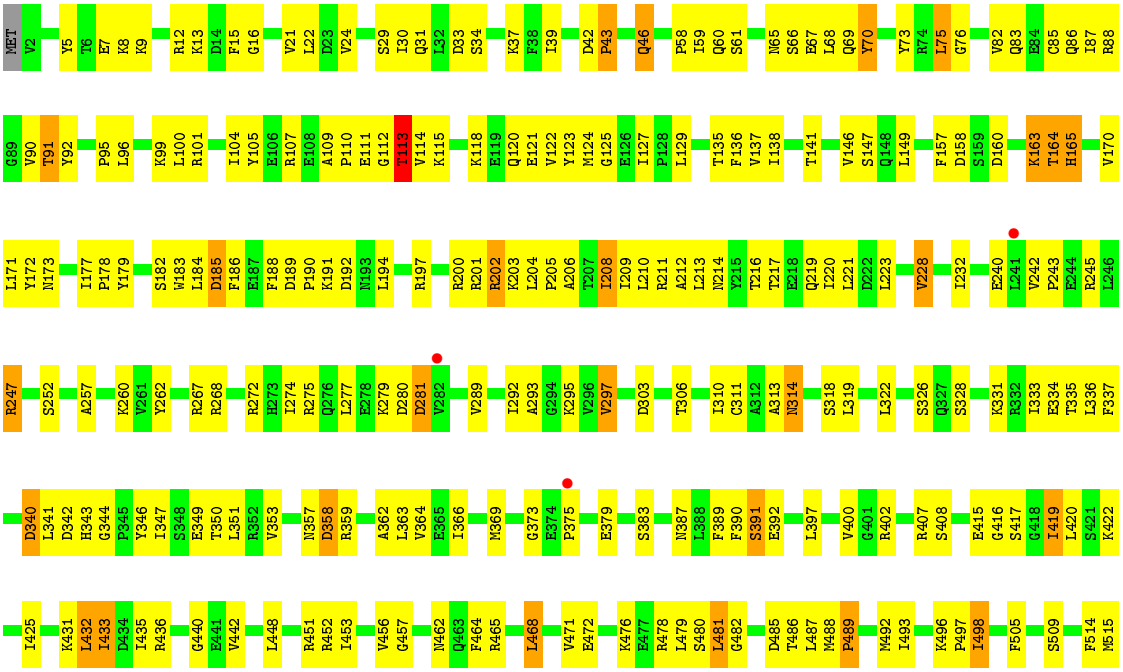
Chain I:

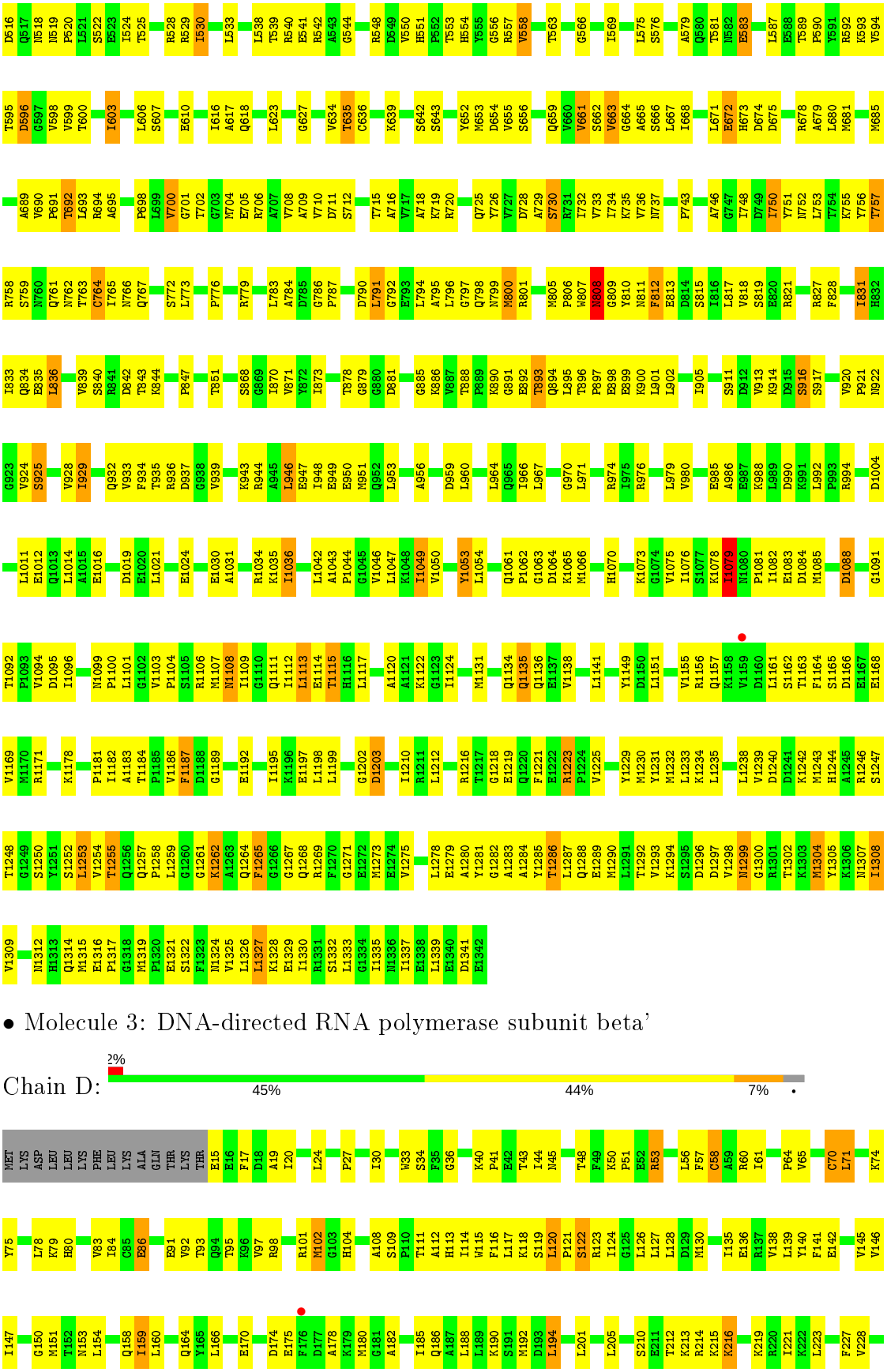




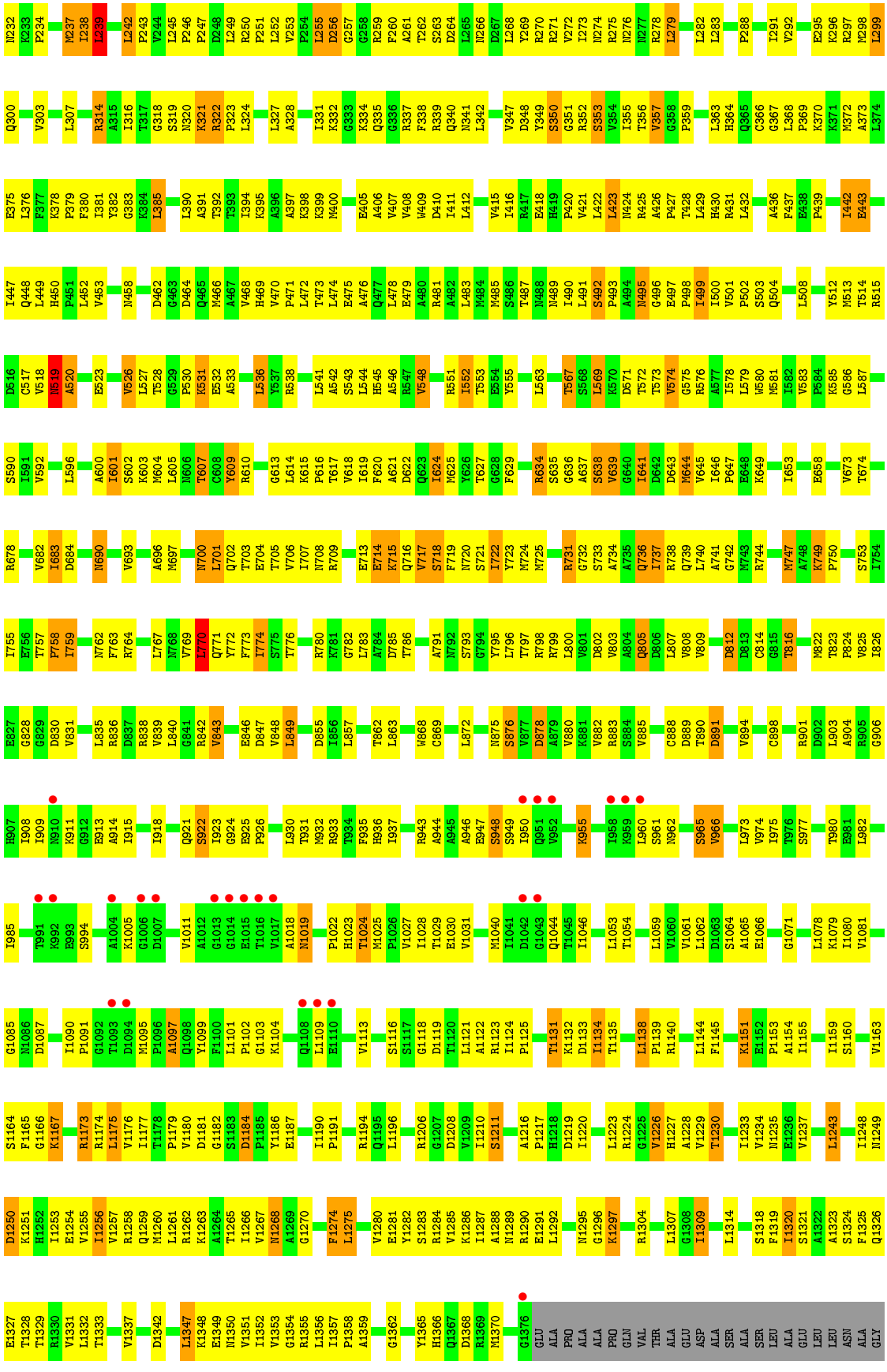
• Molecule 2: DNA-directed RNA polymerase subunit beta

Chain O:





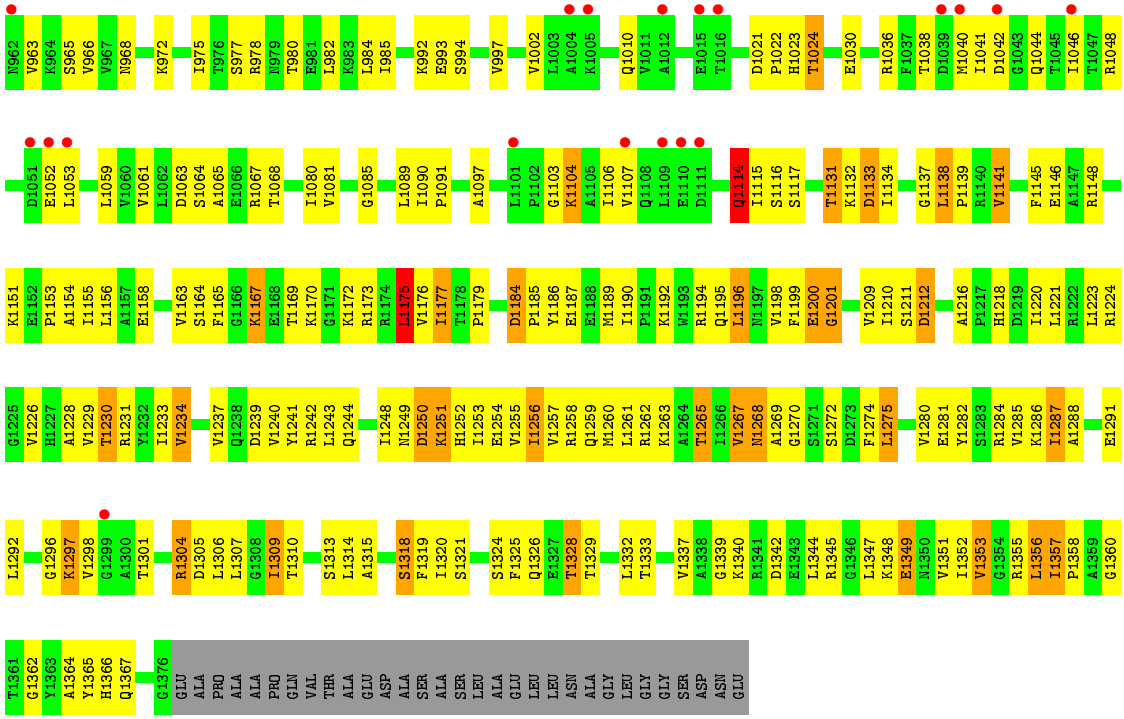
• Molecule 3: DNA-directed RNA polymerase subunit beta'



LEU
GLY
GLY
SER
ASP
ASN
GLU

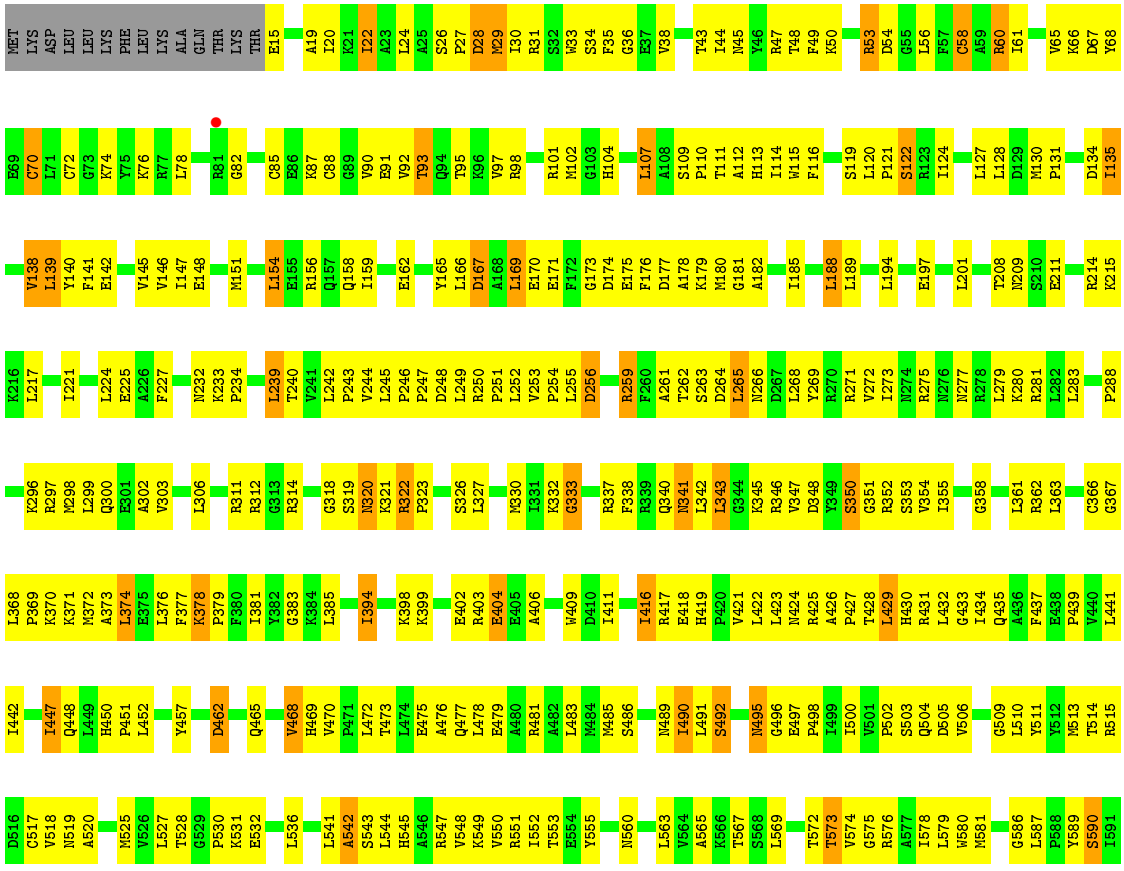
- Molecule 3: DNA-directed RNA polymerase subunit beta'

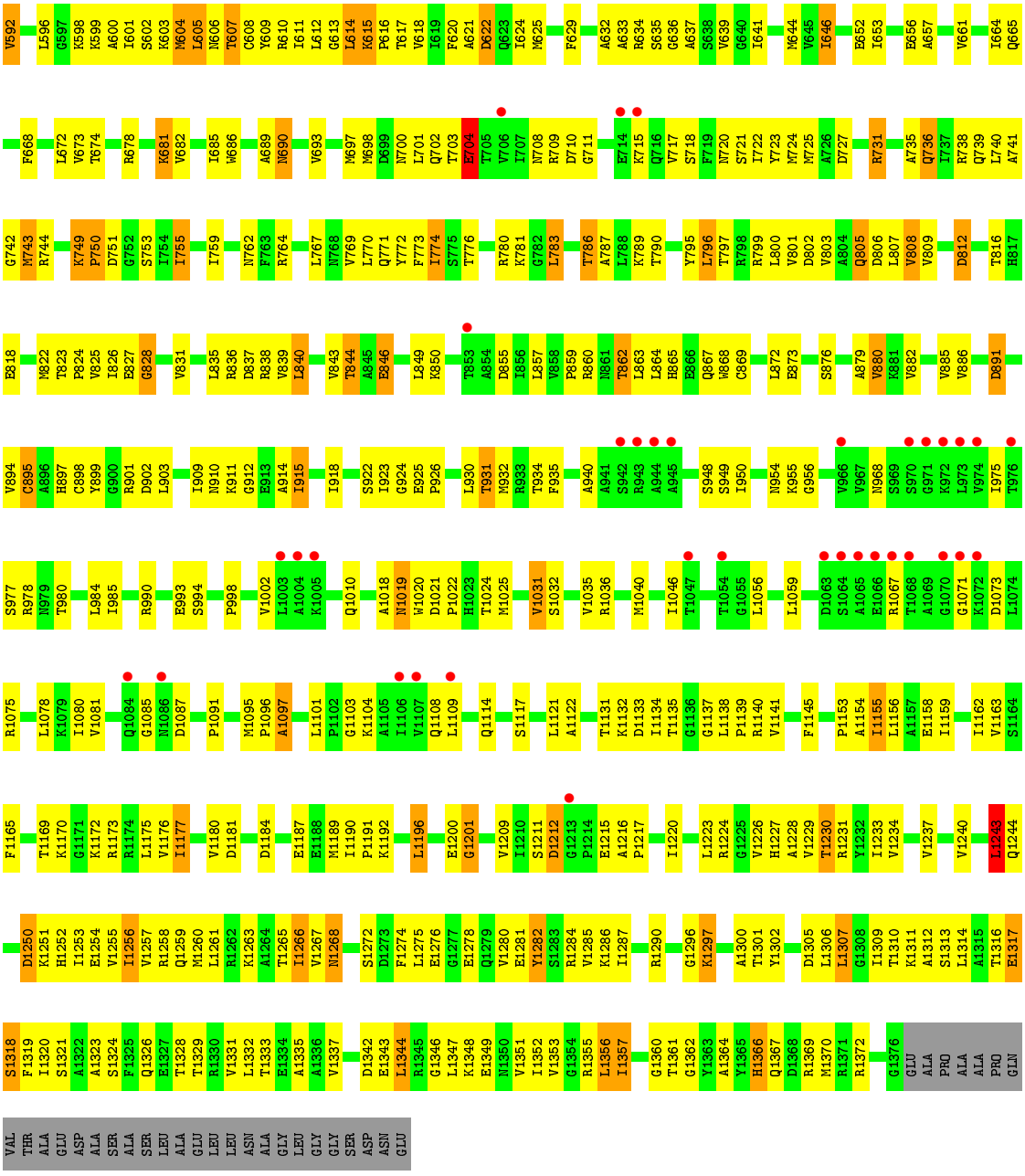
Chain J:



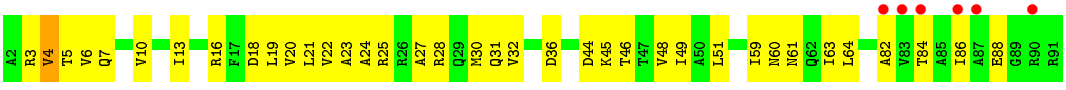
• Molecule 3: DNA-directed RNA polymerase subunit beta'

Chain P:

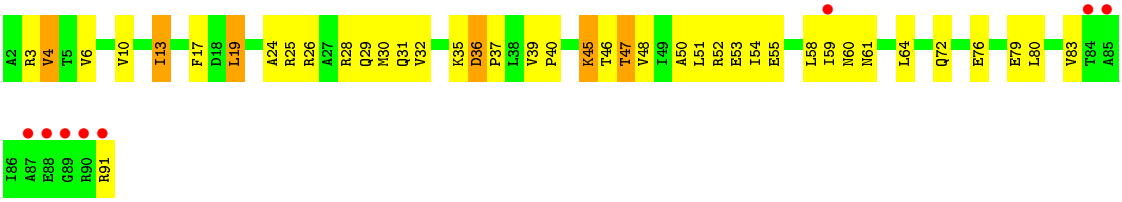




• Molecule 4: DNA-directed RNA polymerase subunit omega

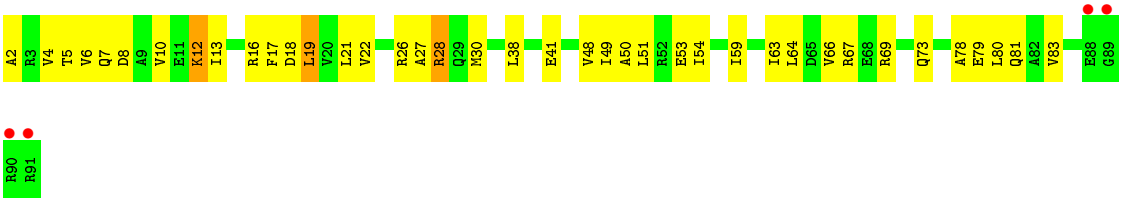


• Molecule 4: DNA-directed RNA polymerase subunit omega



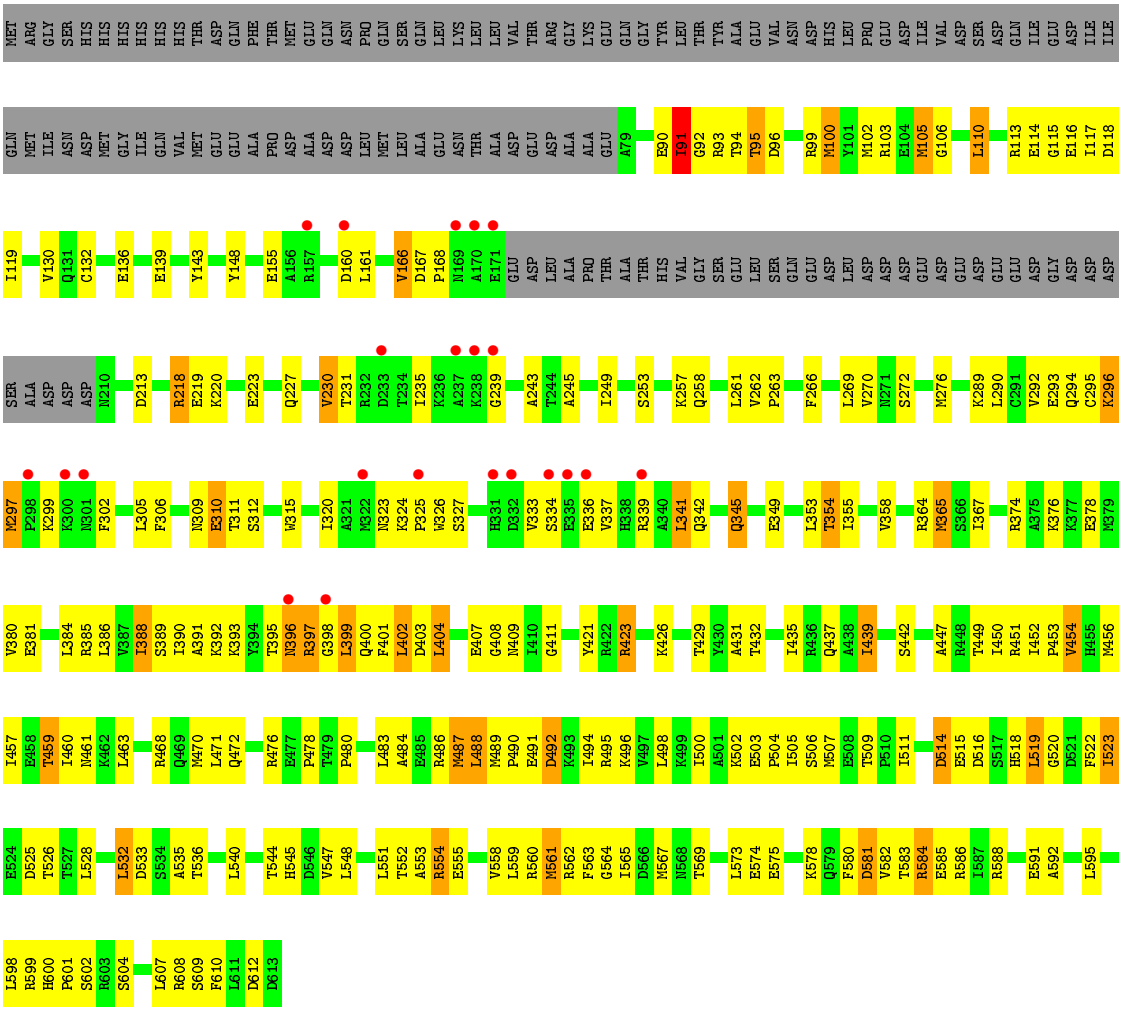
• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain Q:



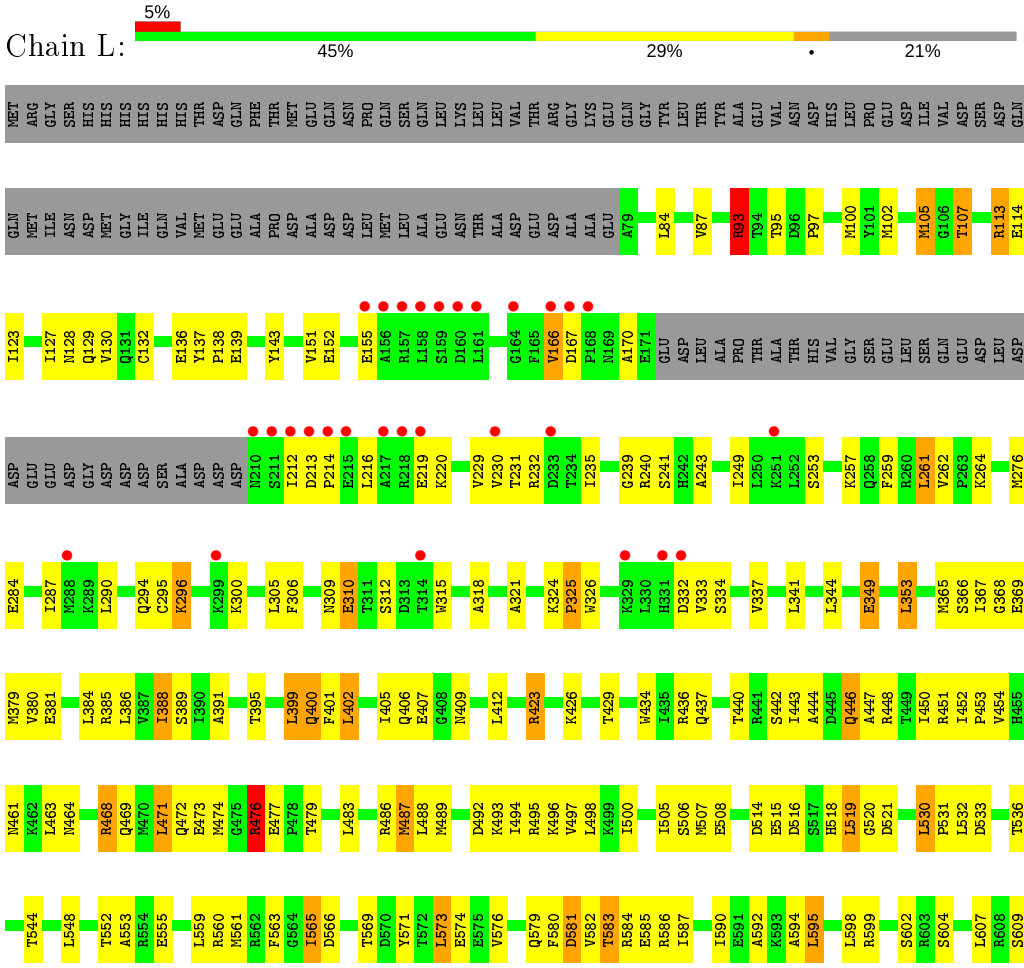
• Molecule 5: RNA polymerase sigma factor RpoD

Chain F:



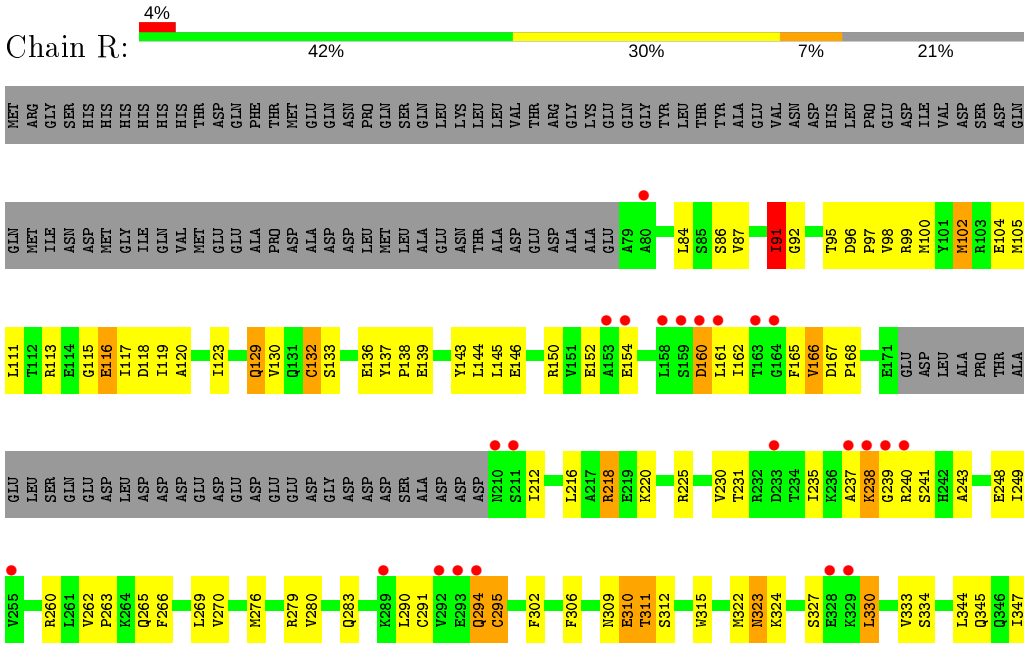
• Molecule 5: RNA polymerase sigma factor RpoD

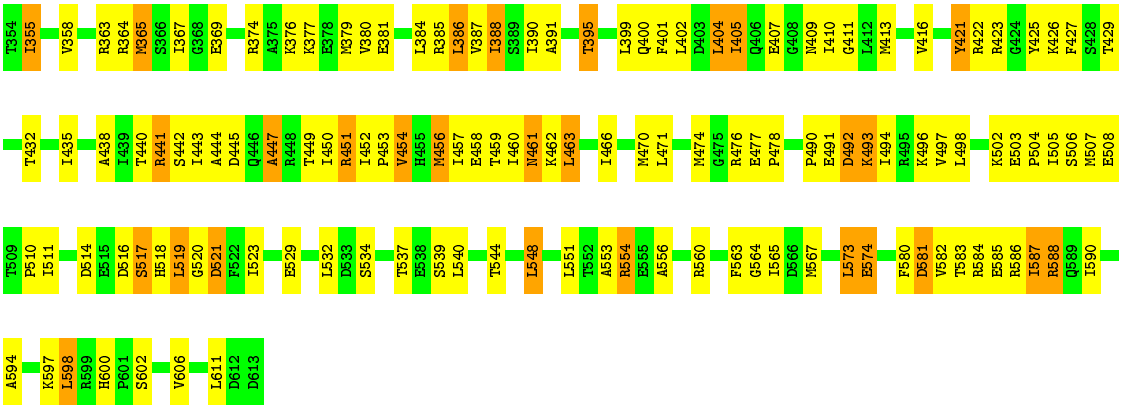
Chain L:



• Molecule 5: RNA polymerase sigma factor RpoD

Chain R:



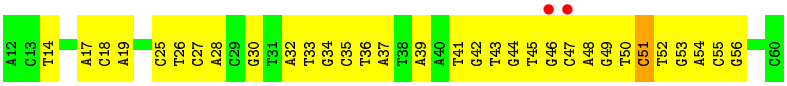


- Molecule 6: NT strand DNA (49-MER)

Chain 1:

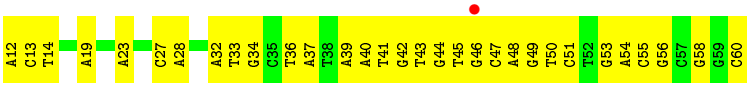
- Molecule 6: NT strand DNA (49-MER)

Chain 4:



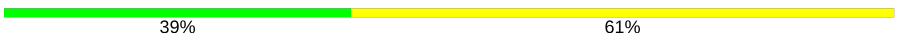
- Molecule 6: NT strand DNA (49-MER)

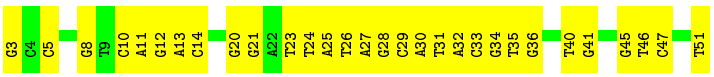
Chain 7:



- Molecule 7: T strand DNA (49-MER)

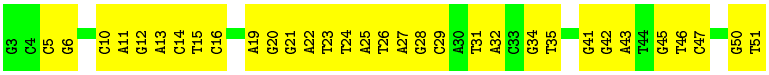
Chain 2:





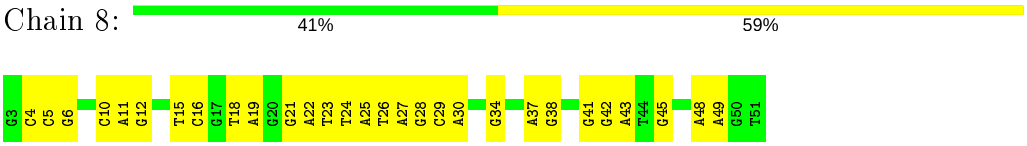
- Molecule 7: T strand DNA (49-MER)

Chain 5:



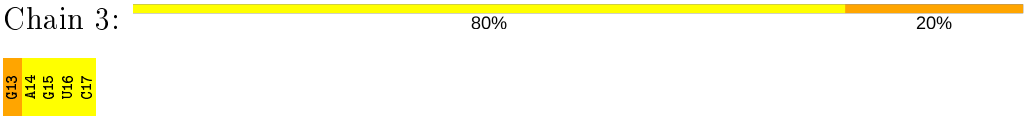
- Molecule 7: T strand DNA (49-MER)

Chain 8:



- Molecule 8: RNA (5'-R(*(GTP))-R(P*AP*GP*UP*C)-3')

Chain 3:

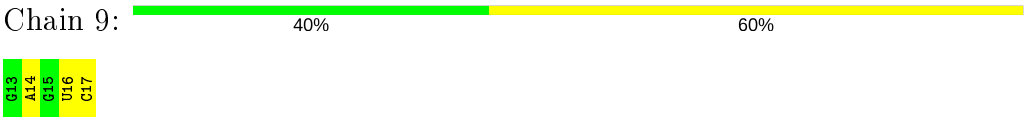


- Molecule 8: RNA (5'-R(*(GTP))-R(P*AP*GP*UP*C)-3')

Chain 6:

- Molecule 8: RNA (5'-R(*(GTP))-R(P*AP*GP*UP*C)-3')

Chain 9:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	237.67Å 204.99Å 248.84Å 90.00° 116.86° 90.00°	Depositor
Resolution (Å)	39.98 – 5.50 39.98 – 5.50	Depositor EDS
% Data completeness (in resolution range)	97.9 (39.98-5.50) 98.1 (39.98-5.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.07 (at 5.37Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.231 , 0.313 0.231 , 0.312	Depositor DCC
R_{free} test set	3384 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	324.1	Xtriage
Anisotropy	0.251	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.24 , 168.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.39$, $< L^2 > = 0.22$	Xtriage
Estimated twinning fraction	0.055 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	94668	wwPDB-VP
Average B, all atoms (Å^2)	198.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.45% 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: GTP, 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	#$\|Z\|>5$	RMSZ	#$\|Z\|>5$
1	A	0.63	0/1809	0.91	5/2450 (0.2%)
1	B	0.58	0/1789	0.87	3/2425 (0.1%)
1	G	0.60	0/1809	0.87	2/2450 (0.1%)
1	H	0.59	0/1789	0.87	2/2425 (0.1%)
1	M	0.53	0/1809	0.76	1/2450 (0.0%)
1	N	0.55	0/1789	0.81	1/2425 (0.0%)
2	C	0.56	0/10745	0.78	5/14499 (0.0%)
2	I	0.58	1/10745 (0.0%)	0.78	5/14499 (0.0%)
2	O	0.53	0/10745	0.75	4/14499 (0.0%)
3	D	0.57	1/10729 (0.0%)	0.80	9/14487 (0.1%)
3	J	0.59	1/10729 (0.0%)	0.85	16/14487 (0.1%)
3	P	0.57	1/10729 (0.0%)	0.80	5/14487 (0.0%)
4	E	0.53	0/710	0.71	0/956
4	K	0.62	1/710 (0.1%)	0.82	0/956
4	Q	0.54	0/710	0.77	0/956
5	F	0.51	0/4076	0.73	1/5482 (0.0%)
5	L	0.53	0/4076	0.75	3/5482 (0.1%)
5	R	0.54	1/4076 (0.0%)	0.75	3/5482 (0.1%)
6	1	0.34	0/1114	0.68	0/1714
6	4	1.27	1/1114 (0.1%)	0.91	4/1714 (0.2%)
6	7	0.40	0/1115	0.66	0/1718
7	2	0.35	0/1136	0.67	0/1752
7	5	0.33	0/1136	0.68	0/1752
7	8	0.41	0/1137	0.66	0/1756
8	3	0.38	0/94	0.67	0/144
8	6	0.42	0/94	0.64	0/144
8	9	0.28	0/94	0.68	0/144
All	All	0.57	7/96608 (0.0%)	0.79	69/131735 (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
3	P	0	1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	4	51	DC	O3'-P	40.58	2.09	1.61
2	I	638	SER	CB-OG	16.07	1.63	1.42
3	D	955	LYS	CE-NZ	10.97	1.76	1.49
4	K	91	ARG	C-O	7.42	1.37	1.23
3	P	681	LYS	CG-CD	5.15	1.70	1.52

The worst 5 of 69 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	4	51	DC	OP1-P-O3'	15.55	139.42	105.20
6	4	51	DC	P-O3'-C3'	15.39	138.17	119.70
6	4	51	DC	O3'-P-O5'	-10.32	84.38	104.00
3	J	120	LEU	C-N-CD	-9.82	99.00	120.60
1	N	29	GLU	C-N-CD	-9.03	100.74	120.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	P	1276	GLU	Peptide

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1787	0	1813	209	0
1	B	1767	0	1789	217	0
1	G	1787	0	1813	166	0
1	H	1767	0	1789	160	0
1	M	1787	0	1813	134	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	N	1767	0	1789	116	0
2	C	10576	0	10591	815	0
2	I	10576	0	10591	916	0
2	O	10576	0	10591	739	0
3	D	10568	0	10781	927	1
3	J	10568	0	10780	1017	0
3	P	10568	0	10783	901	0
4	E	708	0	719	39	0
4	K	708	0	719	38	0
4	Q	708	0	719	47	0
5	F	4022	0	4083	280	0
5	L	4022	0	4083	220	0
5	R	4022	0	4083	298	0
6	1	996	0	555	65	1
6	4	996	0	556	71	0
6	7	996	0	554	60	1
7	2	1012	0	554	55	1
7	5	1012	0	554	53	0
7	8	1012	0	553	48	0
8	3	117	0	55	10	0
8	6	117	0	55	6	0
8	9	117	0	55	6	0
9	D	2	0	0	2	0
9	J	2	0	0	1	0
9	P	2	0	0	5	0
10	D	1	0	0	0	0
10	J	1	0	0	0	0
10	P	1	0	0	0	0
All	All	94668	0	92820	6810	3

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 36.

The worst 5 of 6810 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:I:255:ILE:CG1	2:I:255:ILE:CD1	1.74	1.59
3:D:955:LYS:NZ	3:D:955:LYS:CE	1.76	1.48
3:P:514:THR:HG21	3:P:596:LEU:CD1	1.48	1.42
3:J:421:VAL:CG1	3:J:469:HIS:O	1.70	1.40
3:P:1095:MET:SD	3:P:1173:ARG:NH2	1.97	1.38

All (3) 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)
7:2:3:DG:O5'	7:2:51:DT:O3'[2_657]	1.64	0.56
3:D:1174:ARG:NH2	6:1:17:DA:OP1[2_657]	2.10	0.10
6:7:12:DA:O5'	6:7:60:DC:O3'[2_546]	2.13	0.07

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	228/242 (94%)	213 (93%)	11 (5%)	4 (2%)	8 40
1	B	226/242 (93%)	204 (90%)	14 (6%)	8 (4%)	3 25
1	G	228/242 (94%)	211 (92%)	14 (6%)	3 (1%)	12 47
1	H	226/242 (93%)	205 (91%)	17 (8%)	4 (2%)	8 40
1	M	228/242 (94%)	215 (94%)	12 (5%)	1 (0%)	34 72
1	N	226/242 (93%)	208 (92%)	12 (5%)	6 (3%)	5 31
2	C	1339/1342 (100%)	1220 (91%)	97 (7%)	22 (2%)	9 43
2	I	1339/1342 (100%)	1226 (92%)	88 (7%)	25 (2%)	8 38
2	O	1339/1342 (100%)	1235 (92%)	82 (6%)	22 (2%)	9 43
3	D	1360/1407 (97%)	1212 (89%)	120 (9%)	28 (2%)	7 36
3	J	1360/1407 (97%)	1212 (89%)	113 (8%)	35 (3%)	5 31
3	P	1360/1407 (97%)	1214 (89%)	111 (8%)	35 (3%)	5 31
4	E	88/90 (98%)	84 (96%)	4 (4%)	0	100 100
4	K	88/90 (98%)	84 (96%)	4 (4%)	0	100 100
4	Q	88/90 (98%)	84 (96%)	4 (4%)	0	100 100
5	F	493/628 (78%)	449 (91%)	30 (6%)	14 (3%)	5 30
5	L	493/628 (78%)	444 (90%)	30 (6%)	19 (4%)	3 23

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
5	R	493/628 (78%)	447 (91%)	30 (6%)	16 (3%)	4 26
All	All	11202/11853 (94%)	10167 (91%)	793 (7%)	242 (2%)	6 35

5 of 242 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	210	THR
1	B	209	GLY
2	C	165	HIS
2	C	808	ASN
2	C	812	PHE

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	198/208 (95%)	166 (84%)	32 (16%)	2 14
1	B	196/208 (94%)	163 (83%)	33 (17%)	2 12
1	G	198/208 (95%)	180 (91%)	18 (9%)	9 31
1	H	196/208 (94%)	171 (87%)	25 (13%)	4 19
1	M	198/208 (95%)	183 (92%)	15 (8%)	13 39
1	N	196/208 (94%)	179 (91%)	17 (9%)	10 33
2	C	1156/1157 (100%)	1027 (89%)	129 (11%)	6 23
2	I	1156/1157 (100%)	1038 (90%)	118 (10%)	7 26
2	O	1156/1157 (100%)	1044 (90%)	112 (10%)	8 28
3	D	1135/1168 (97%)	1009 (89%)	126 (11%)	6 24
3	J	1135/1168 (97%)	1003 (88%)	132 (12%)	5 22
3	P	1135/1168 (97%)	1014 (89%)	121 (11%)	6 25
4	E	74/74 (100%)	71 (96%)	3 (4%)	30 55
4	K	74/74 (100%)	65 (88%)	9 (12%)	5 21
4	Q	74/74 (100%)	68 (92%)	6 (8%)	11 37

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
5	F	439/554 (79%)	395 (90%)	44 (10%)	7 27
5	L	439/554 (79%)	401 (91%)	38 (9%)	10 33
5	R	439/554 (79%)	384 (88%)	55 (12%)	4 20
All	All	9594/10107 (95%)	8561 (89%)	1033 (11%)	6 25

5 of 1033 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	I	734	ILE
3	J	567	THR
3	P	1177	ILE
2	I	836	LEU
2	I	1337	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 143 such sidechains are listed below:

Mol	Chain	Res	Type
2	I	1116	HIS
3	J	690	ASN
3	P	1114	GLN
2	I	1307	ASN
3	J	341	ASN

5.3.3 RNA

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	3	4/5 (80%)	0	1 (25%)
8	6	4/5 (80%)	0	1 (25%)
8	9	3/5 (60%)	0	0
All	All	11/15 (73%)	0	2 (18%)

There are no RNA backbone outliers to report.

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
8	3	13	GTP
8	6	13	GTP

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

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
6	4	2
7	2	1
6	1	1
7	5	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	4	46:DG	O3'	47:DC	P	5.33
1	1	46:DG	O3'	47:DC	P	4.95
1	2	12:DG	O3'	13:DA	P	2.74
1	5	11:DA	O3'	12:DG	P	2.33

Continued on next page...

Continued from previous page...

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	4	51:DC	O3'	52:DT	P	2.09

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	230/242 (95%)	-0.42	1 (0%) 92 87	134, 152, 183, 205	0
1	B	228/242 (94%)	-0.55	0 100 100	136, 167, 199, 236	0
1	G	230/242 (95%)	-0.30	1 (0%) 92 87	139, 162, 198, 240	0
1	H	228/242 (94%)	-0.40	3 (1%) 77 68	141, 176, 208, 242	0
1	M	230/242 (95%)	-0.20	4 (1%) 70 61	159, 179, 209, 245	0
1	N	228/242 (94%)	-0.15	6 (2%) 56 47	169, 201, 249, 272	0
2	C	1341/1342 (99%)	-0.34	3 (0%) 95 93	107, 166, 250, 351	0
2	I	1341/1342 (99%)	-0.36	3 (0%) 95 93	98, 172, 227, 283	0
2	O	1341/1342 (99%)	-0.35	4 (0%) 94 90	113, 174, 222, 263	0
3	D	1362/1407 (96%)	-0.21	26 (1%) 66 58	112, 184, 269, 324	0
3	J	1362/1407 (96%)	-0.22	26 (1%) 66 58	100, 172, 323, 386	0
3	P	1362/1407 (96%)	-0.17	36 (2%) 56 47	117, 182, 291, 333	0
4	E	90/90 (100%)	0.07	6 (6%) 17 16	136, 169, 350, 413	0
4	K	90/90 (100%)	-0.10	8 (8%) 9 11	112, 152, 324, 363	0
4	Q	90/90 (100%)	-0.30	4 (4%) 34 30	128, 171, 328, 364	0
5	F	497/628 (79%)	-0.11	22 (4%) 34 30	154, 271, 387, 434	0
5	L	497/628 (79%)	0.07	29 (5%) 23 21	138, 281, 365, 402	0
5	R	497/628 (79%)	-0.13	23 (4%) 32 29	146, 261, 390, 426	0
6	1	49/49 (100%)	-0.34	0 100 100	205, 265, 288, 289	0
6	4	49/49 (100%)	-0.34	2 (4%) 37 32	181, 228, 278, 302	0
6	7	49/49 (100%)	-0.38	1 (2%) 65 57	184, 228, 266, 277	0
7	2	49/49 (100%)	-0.51	0 100 100	192, 268, 291, 312	0
7	5	49/49 (100%)	-0.29	0 100 100	163, 232, 279, 326	0
7	8	49/49 (100%)	-0.51	0 100 100	166, 227, 262, 322	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
8	3	4/5 (80%)	0.18	0	100	100	230, 234, 236, 245	0
8	6	4/5 (80%)	0.03	0	100	100	220, 221, 224, 239	0
8	9	4/5 (80%)	0.33	0	100	100	215, 221, 224, 236	0
All	All	11550/12162 (94%)	-0.25	208 (1%)	68	60	98, 182, 331, 434	0

The worst 5 of 208 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	L	211	SER	11.1
3	D	959	LYS	7.7
5	L	212	ILE	6.8
5	R	211	SER	6.7
3	P	1004	ALA	6.5

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

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

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. 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
9	ZN	P	1501	1/1	0.92	0.08	206,206,206,206	0
9	ZN	D	1501	1/1	0.93	0.06	220,220,220,220	0
9	ZN	J	1501	1/1	0.94	0.07	211,211,211,211	0
9	ZN	P	1502	1/1	0.96	0.14	158,158,158,158	0
9	ZN	J	1502	1/1	0.96	0.17	144,144,144,144	0
9	ZN	D	1502	1/1	0.98	0.15	181,181,181,181	0
10	MG	P	1503	1/1	0.98	0.31	170,170,170,170	0
10	MG	J	1503	1/1	0.99	0.19	145,145,145,145	0
10	MG	D	1503	1/1	0.99	0.16	141,141,141,141	0

6.5 Other polymers (i)

There are no such residues in this entry.