wwPDB X-ray Structure Validation Summary Report (i)

Aug 16, 2023 – 03:54 PM EDT

PDB ID	:	1ZYR
Title	:	Structure of Thermus thermophilus RNA polymerase holoenzyme in complex with the antibiotic streptolydigin
Authors	:	Tuske, S.; Sarafianos, S.G.; Wang, X.; Hudson, B.; Sineva, E.; Mukhopadhyay, J.; Birktoft, J.J.; Leroy, O.; Ismail, S.; Clark, A.D.; Dharia, C.; Napoli, A.; Laptenko, O.; Lee, J.; Borukhov, S.; Ebright, R.H.; Arnold, E.
Deposited on	:	2005-06-10
Resolution	:	3.00 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

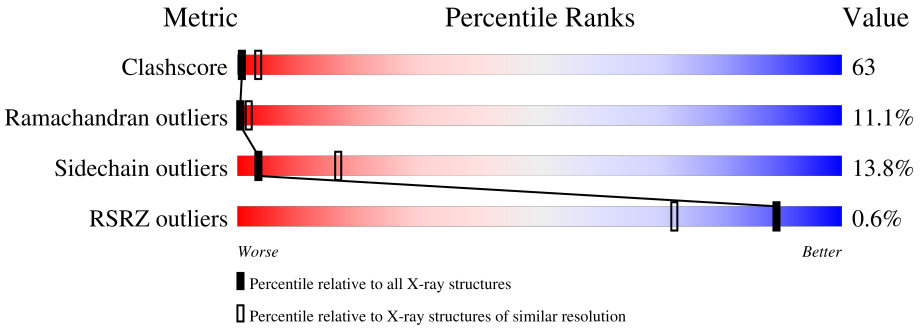
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



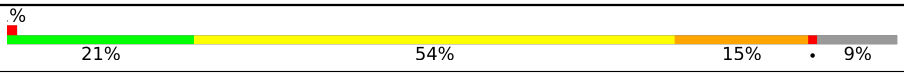
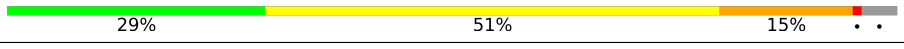
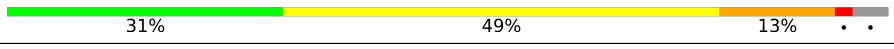
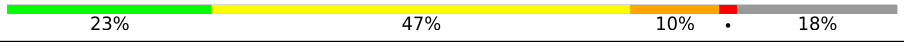
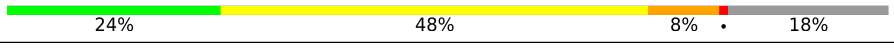
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	315	15% 45% 11% • 27%
1	B	315	21% 43% 8% 27%
1	K	315	16% 44% 11% • 27%
1	L	315	21% 41% 11% 27%
2	C	1119	25% 59% 14% •
2	M	1119	25% 59% 14% •
3	D	1524	% 20% 53% 17% • 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	N	1524	
4	E	99	
4	O	99	
5	F	423	
5	P	423	

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 54048 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 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
1	B	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
1	K	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			
1	L	229	Total	C	N	O	S	0	0	0
			1806	1153	313	337	3			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1119	Total	C	N	O	S	0	0	0
			8829	5581	1577	1647	24			
2	M	1119	Total	C	N	O	S	0	0	0
			8829	5581	1577	1647	24			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1392	Total	C	N	O	S	0	0	0
			10975	6953	1941	2048	33			
3	N	1392	Total	C	N	O	S	0	0	0
			10975	6953	1941	2048	33			

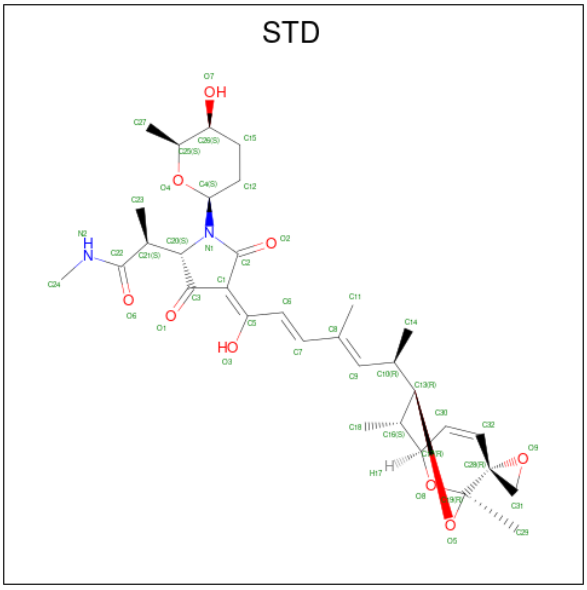
- Molecule 4 is a protein called DNA-directed RNA polymerase omega chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	95	Total	C	N	O	S	0	0	0
			769	488	133	144	4			
4	O	95	Total	C	N	O	S	0	0	0
			769	488	133	144	4			

- Molecule 5 is a protein called DNA-directed RNA polymerase sigma chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	345	Total	C	N	O	S	0	0	0
			2793	1762	504	523	4			
5	P	345	Total	C	N	O	S	0	0	0
			2793	1762	504	523	4			

- Molecule 6 is STREPTOLYDIGIN (three-letter code: STD) (formula: $C_{32}H_{44}N_2O_9$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	D	1	Total	C	N	O	0	0
			43	32	2	9		
6	M	1	Total	C	N	O	0	0
			43	32	2	9		

- Molecule 7 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	D	2	Total	Zn	0	0
			2	2		
7	N	2	Total	Zn	0	0
			2	2		

- Molecule 8 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

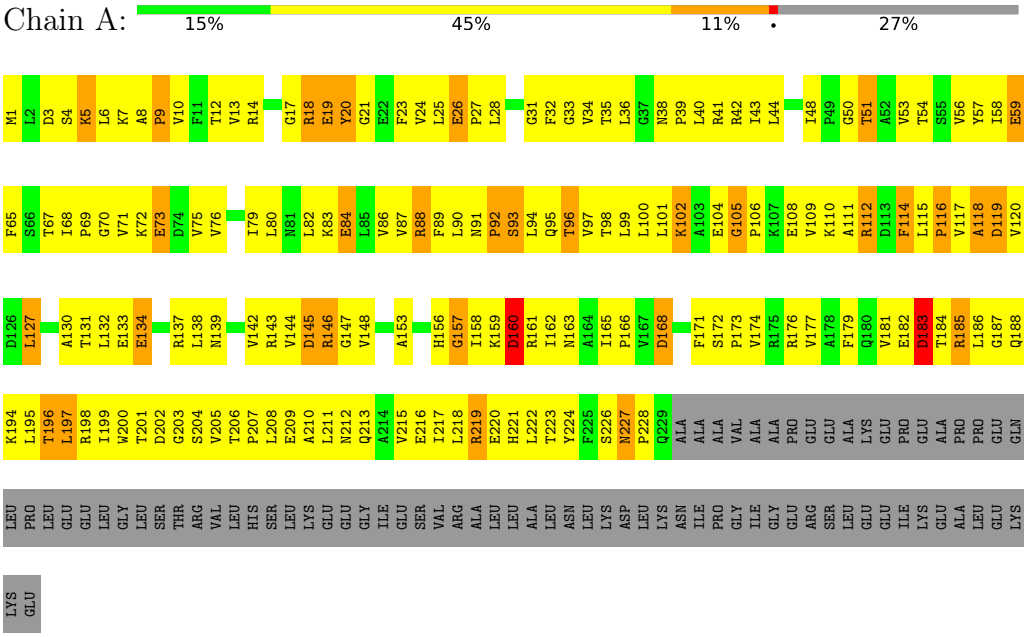
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	D	1	Total Mg 1 1	0	0
8	N	1	Total Mg 1 1	0	0

3 Residue-property plots (i)

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

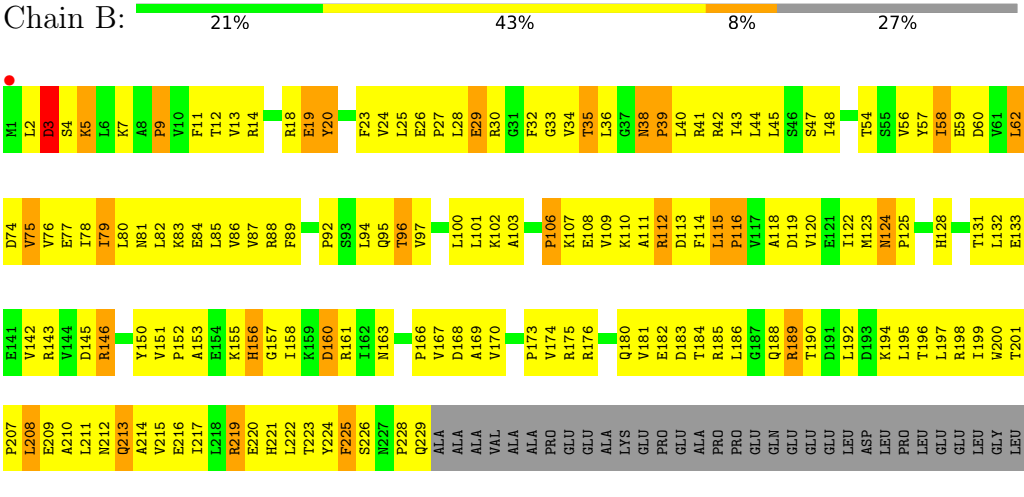
• Molecule 1: DNA-directed RNA polymerase alpha chain

Chain A:



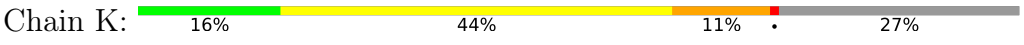
• Molecule 1: DNA-directed RNA polymerase alpha chain

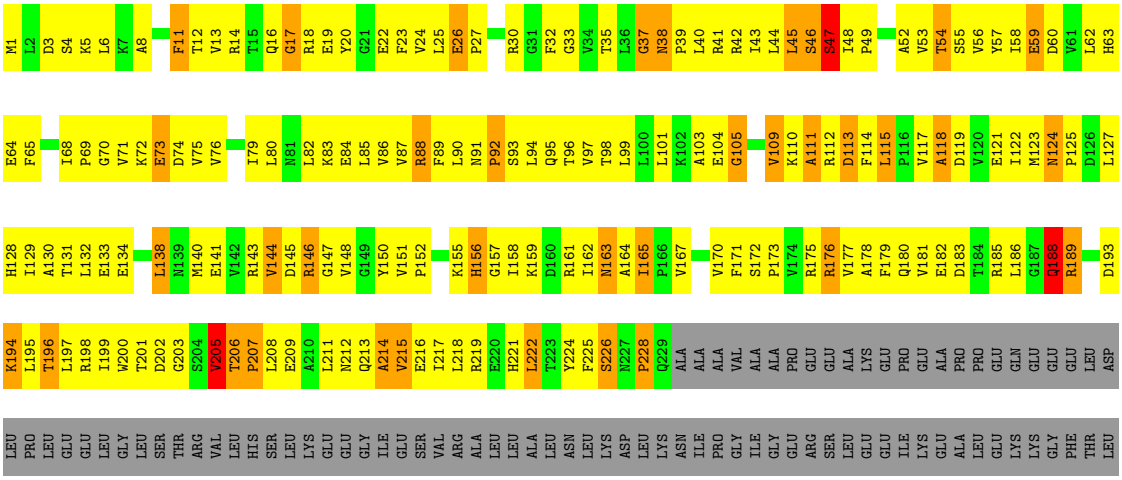
Chain B:



• Molecule 1: DNA-directed RNA polymerase alpha chain

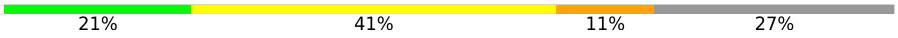
Chain K:

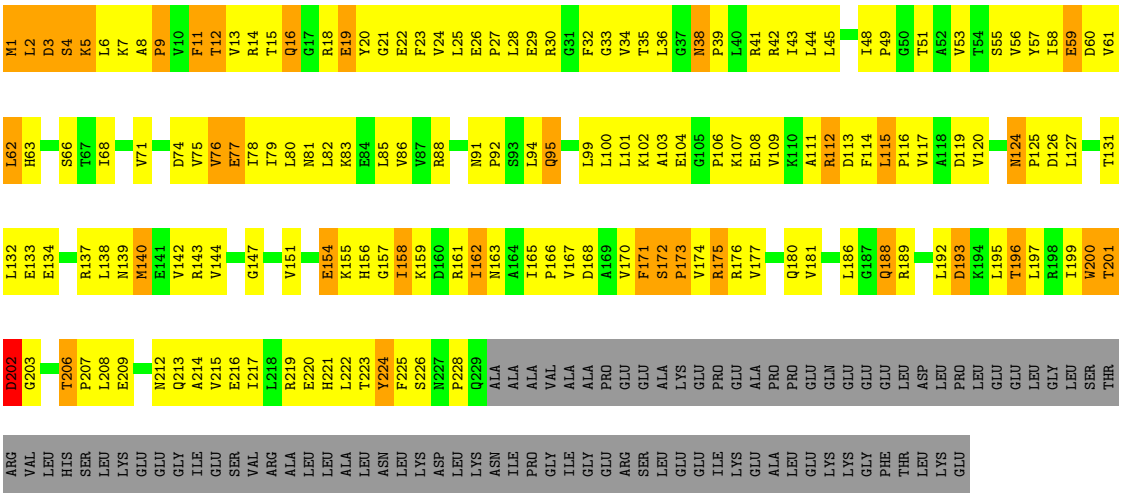




• Molecule 1: DNA-directed RNA polymerase alpha chain

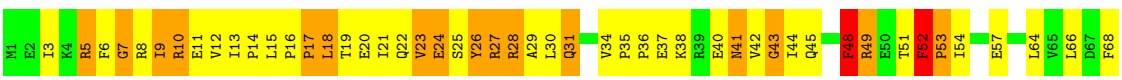
Chain L:

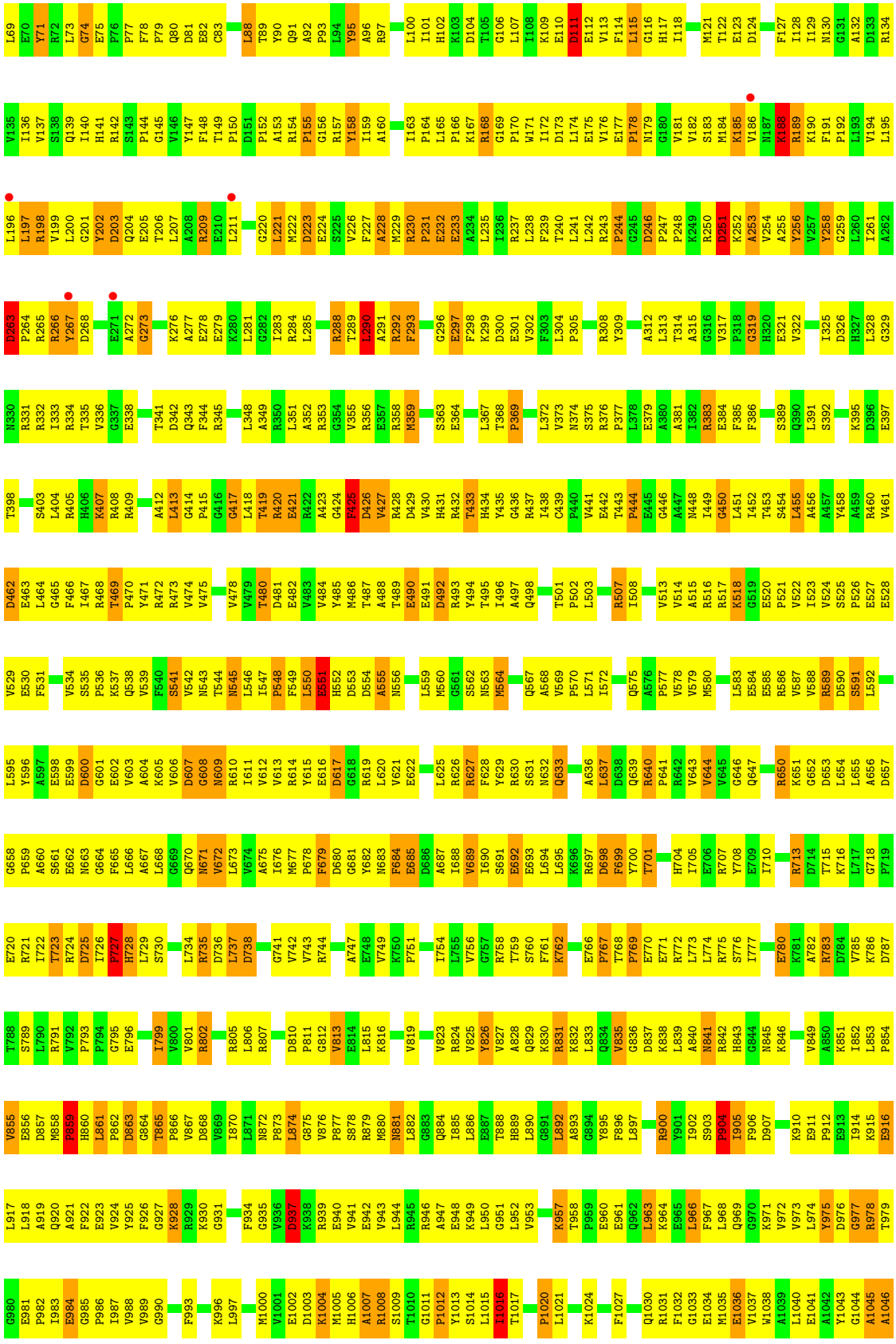




• Molecule 2: DNA-directed RNA polymerase beta chain

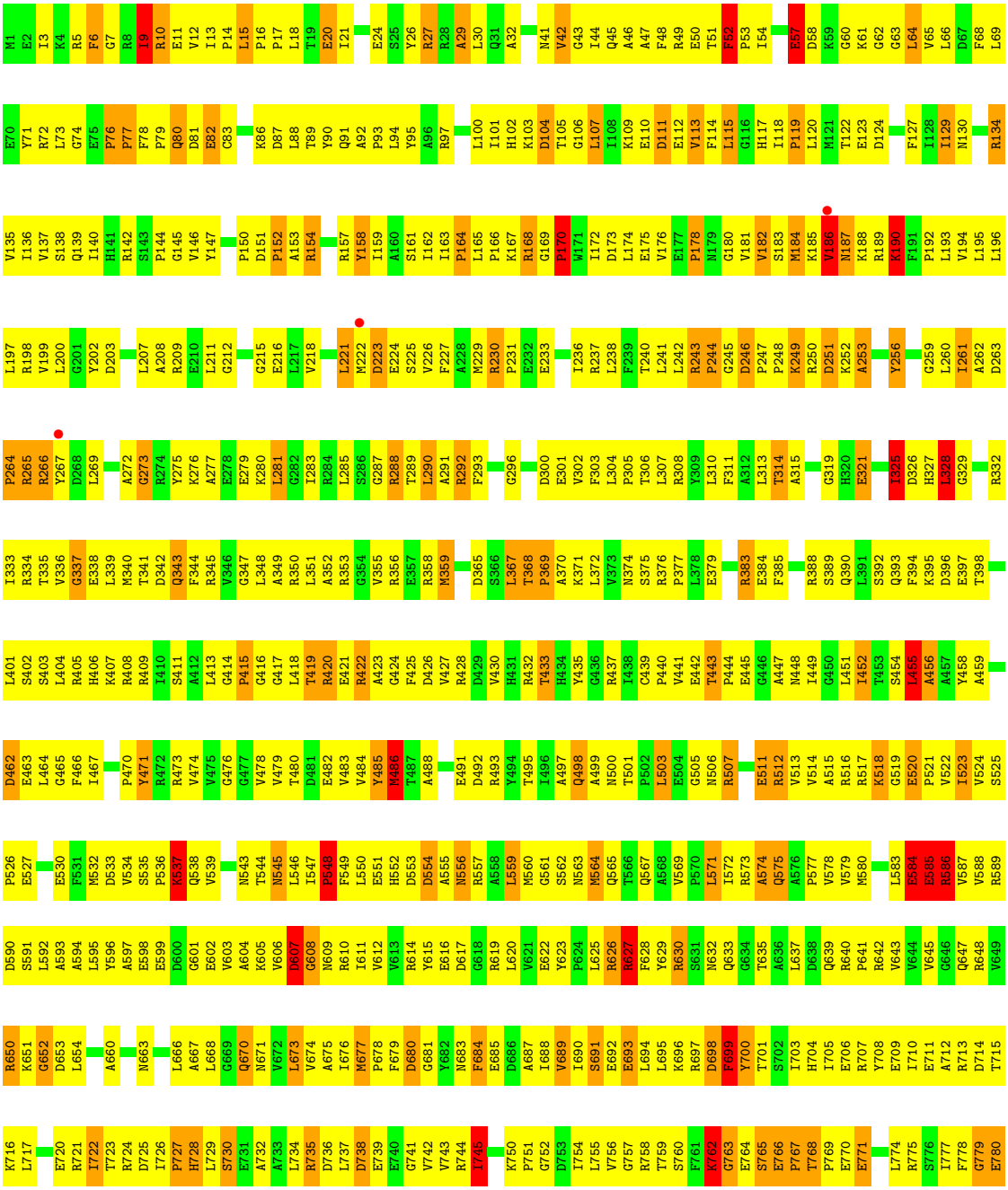
Chain C:

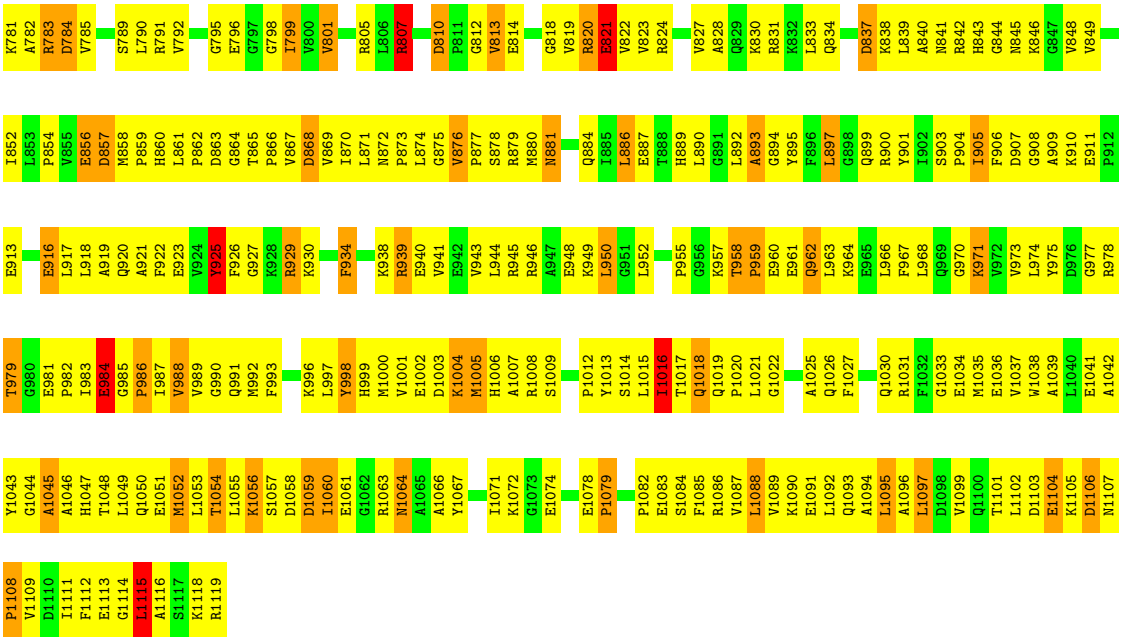



L69 E70 Y71 R72 L73 G74 E75 P76 P77 F78 P79 Q80 D81 E82 C83 L88 T89 Y90 Q91 A92 P93 L94 Y95 A96 R97 L100 I101 H102 K103 D104 T105 G106 L107 I108 K109 E110 D111 E112 V113 F114 L115 G116 H117 I118 M121 T122 E123 D124 F127 I128 I129 N130 G131 A132 D133 R134
V135 I136 V137 S138 Q139 I140 H141 R142 S143 P144 G145 V146 Y147 F148 T149 P150 D151 P152 A153 R154 P155 G156 R157 Y158 I159 A160 I163 P164 L165 P166 K167 R168 G169 P170 W171 I172 D173 L174 E175 V176 E177 P178 N179 G180 V181 V182 S183 M184 K185 V186 N187 K188 R189 K190 F191 P192 L193 V194 L195
L196 L197 R198 V199 L200 G201 Y202 D203 Q204 E205 T206 L207 A208 R209 E210 L211 G220 L221 M222 D223 E224 S225 V226 F227 A228 M229 R230 P231 E232 E233 A234 L235 I236 R237 L238 F239 T240 L241 L242 R243 P244 G245 D246 P247 P248 K249 R250 D251 K252 A253 V254 A255 Y256 V257 Y258 G259 L260 I261 A262
D263 P264 R265 R266 Y267 D268 E271 A272 G273 K276 A277 E278 E279 K280 L281 G282 I283 R284 L285 R288 T289 L290 A291 R292 F293 G296 E297 F298 K299 D300 E301 V302 F303 L304 P305 R308 Y309 A312 L313 T314 A315 G316 V317 P318 G319 H320 E321 V322 I325 D326 H327 L328 G329
N330 R331 R332 I333 R334 T335 V336 G337 E338 T341 D342 Q343 F344 R345 L348 A349 R350 L351 A352 R353 G354 V355 R356 E357 R358 M359 S363 E364 L367 T368 P369 L372 V373 N374 S375 R376 P377 L378 E379 A380 A381 I382 R383 E384 F385 F386 S389 Q390 L391 S392 K395 D396 E397
T398 S403 L404 R405 H406 K407 R408 R409 A412 L413 G414 P415 G416 G417 L418 T419 R420 E421 R422 A423 G424 F425 D426 V427 R428 D429 V430 H431 R432 T433 H434 Y435 G436 R437 I438 C439 P440 V441 E442 T443 P444 E445 G446 A447 N448 I449 G450 L451 I452 T453 S454 L455 A456 A457 Y458 A459 R460 V461
D462 E463 L464 G465 F466 I467 R468 T469 P470 Y471 R472 R473 V474 V475 V478 V479 T480 D481 E482 V483 V484 Y485 M486 T487 A488 T489 E490 E491 D492 R493 Y494 T495 I496 A497 Q498 T501 P502 L503 R507 I508 V513 V514 A515 R516 R517 K518 G519 E520 P521 V522 I523 V524 S525 P526 E527 E528
V529 E530 F531 V534 S535 P536 K537 Q538 V539 F540 S541 V542 N543 T544 N545 L546 I547 P548 F549 L550 E551 H552 D553 D554 A555 N556 L559 M560 G561 S562 N563 M564 Q567 A568 V569 P570 L571 I572 Q575 A576 P577 V578 V579 M580 L583 E584 E585 R586 V587 V588 R589 D590 S591 L592
L595 Y596 A597 E598 E599 D600 G601 E602 V603 A604 K605 V606 D607 G608 N609 R610 I611 V612 V613 R614 Y615 E616 D617 G618 R619 L620 V621 E622 L625 R626 R627 F628 Y629 R630 S631 N632 Q633 A636 L637 D638 Q639 R640 P641 R642 V643 V644 V645 G646 Q647 R650 K651 G652 D653 L654 L655 A656 D657
G658 P659 A660 S661 E662 N663 G664 F665 L666 A667 L668 G669 Q670 N671 V672 L673 V674 A675 I676 M677 P678 F679 D680 G681 Y682 N683 F684 E685 D686 A687 I688 V689 I690 S691 E692 E693 L694 L695 K696 R697 D698 F699 Y700 T701 H704 I705 E706 R707 Y708 E709 I710 R713 D714 T715 K716 L717 G718 P719
E720 R721 I722 T723 R724 D725 I726 P727 H728 L729 S730 L734 R735 D736 L737 D738 G741 V742 V743 R744 A747 E748 V749 K750 P751 I754 L755 V756 G757 R758 T759 S760 F761 K762 E766 P767 T768 P769 E770 E771 R772 L773 L774 R775 S776 I777 E780 K781 A782 R783 D784 V785 K786 D787
T788 S789 L790 R791 V792 P793 P794 G795 E796 I799 V800 V801 R802 R805 L806 R807 D810 P811 G812 V813 E814 L815 K816 V819 V823 R824 V825 Y826 V827 A828 Q829 K830 R831 K832 L833 Q834 V835 G836 D837 K838 L839 A840 N841 R842 H843 G844 N845 K846 V849 A850 K851 I852 I853 P854
V855 E856 D857 M858 P859 H860 L861 P862 D863 G864 T865 P866 V867 D868 V869 I870 L871 N872 P873 L874 G875 V876 P877 S878 R879 M880 N881 L882 G883 Q884 I885 L886 E887 T888 H889 L890 G891 L892 A893 G894 Y895 F896 L897 R900 Y901 I902 S903 P904 I905 F906 D907 K910 E911 P912 E913 I914 K915 E916
L917 L918 A919 Q920 A921 F922 E923 V924 Y925 F926 G927 K928 R929 K930 G931 F934 G935 V936 D937 K938 R939 E940 V941 E942 V943 L944 R945 R946 A947 E948 K949 L950 G951 L952 V953 K957 T958 P959 E960 E961 Q962 L963 K964 E965 L966 F967 L968 Q969 G970 K971 V972 V973 L974 Y975 D976 G977 R978 T979
G980 E981 P982 I983 E984 G985 P986 I987 V988 V989 G990 F993 K996 L997 M1000 V1001 E1002 D1003 K1004 M1005 H1006 A1007 R1008 S1009 T1010 G1011 P1012 Y1013 S1014 L1015 I1016 T1017 P1020 L1021 K1024 F1027 Q1030 R1031 F1032 G1033 E1034 M1035 E1036 V1037 W1038 A1039 L1040 E1041 A1042 Y1043 G1044 A1045 A1046

• Molecule 2: DNA-directed RNA polymerase beta chain

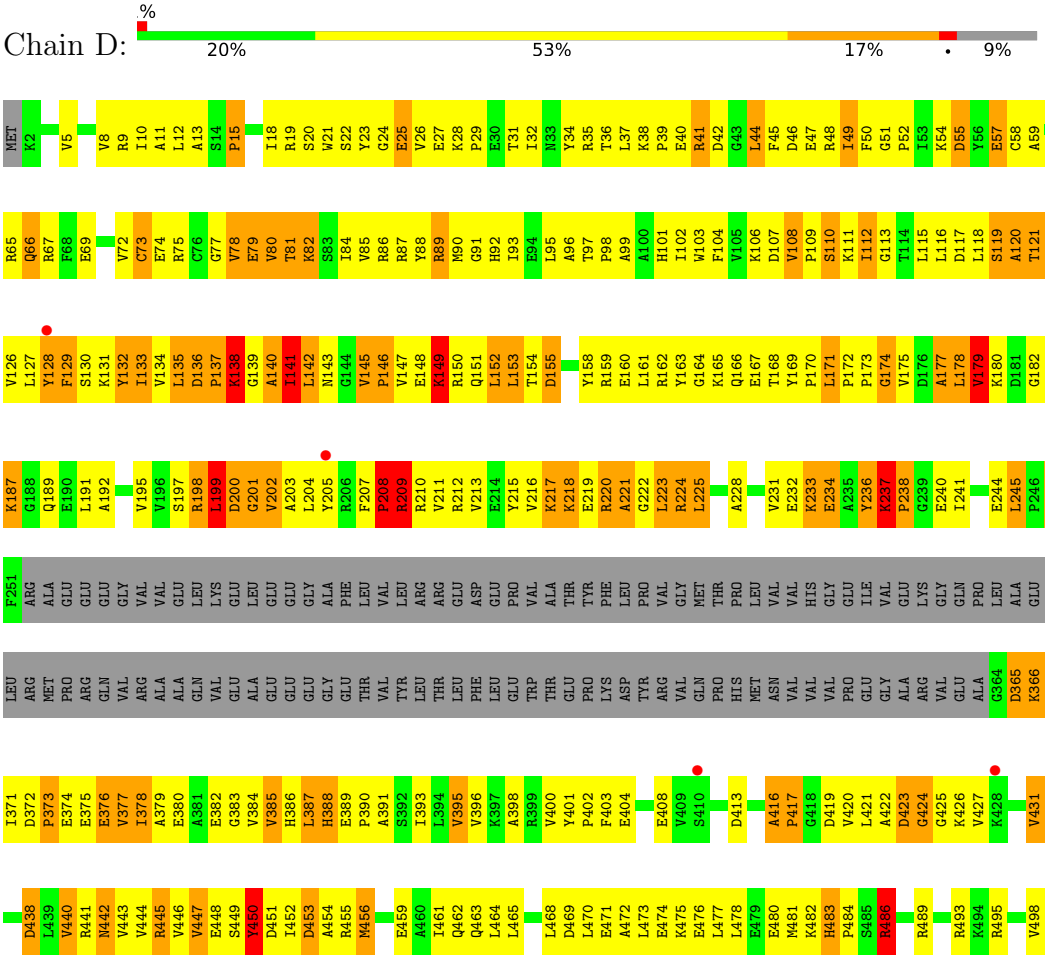
Chain M:

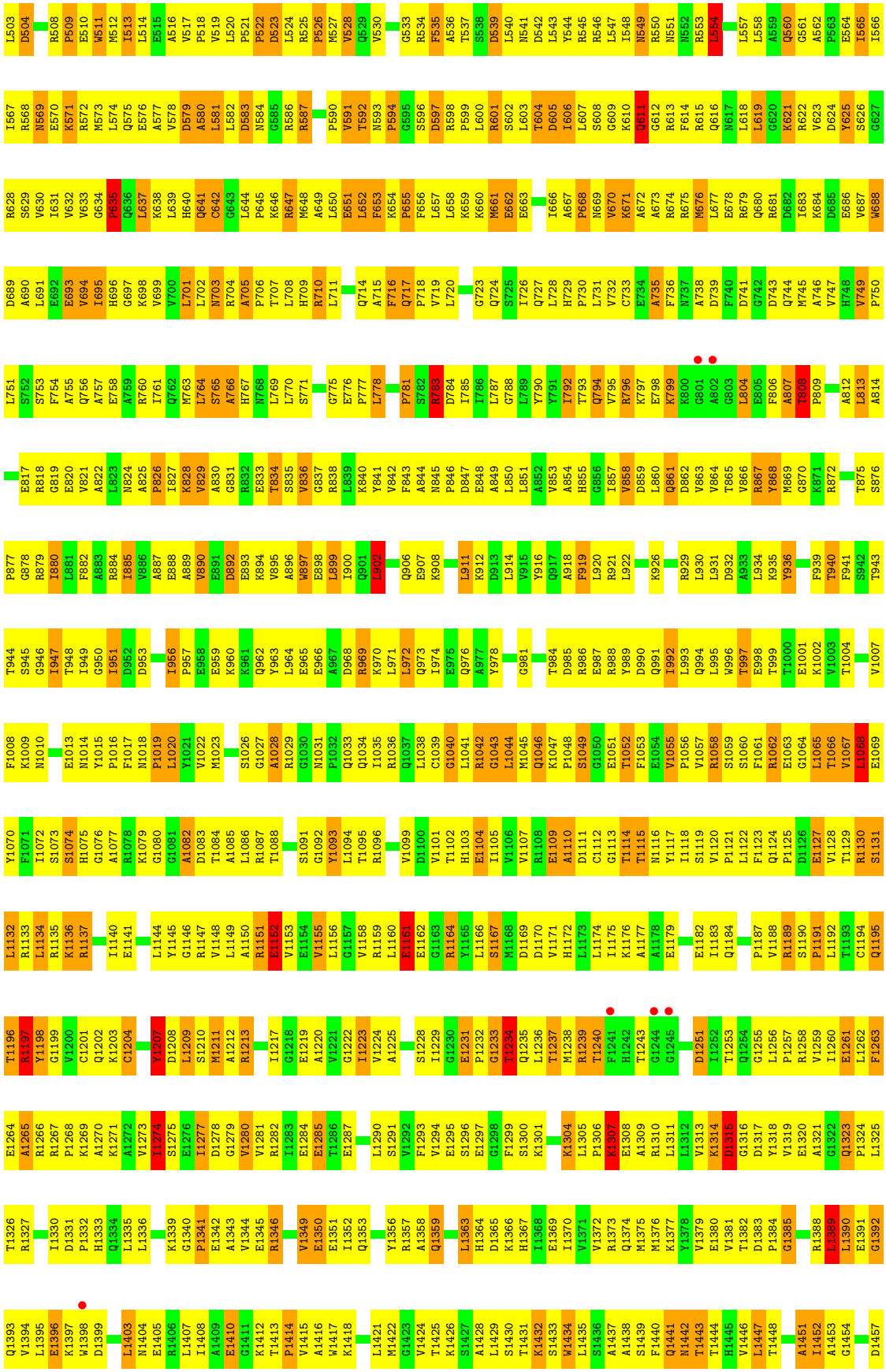


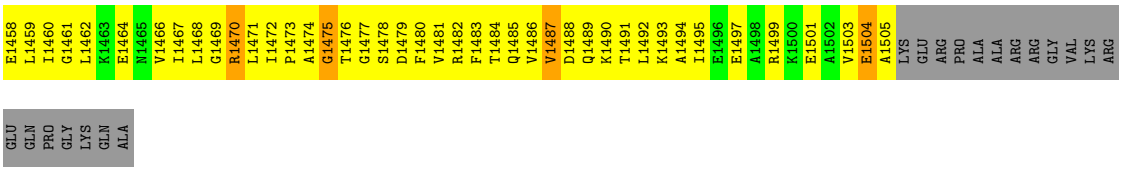


• Molecule 3: DNA-directed RNA polymerase subunit beta' chain

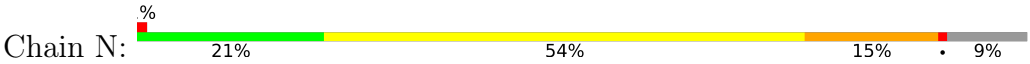
Chain D:

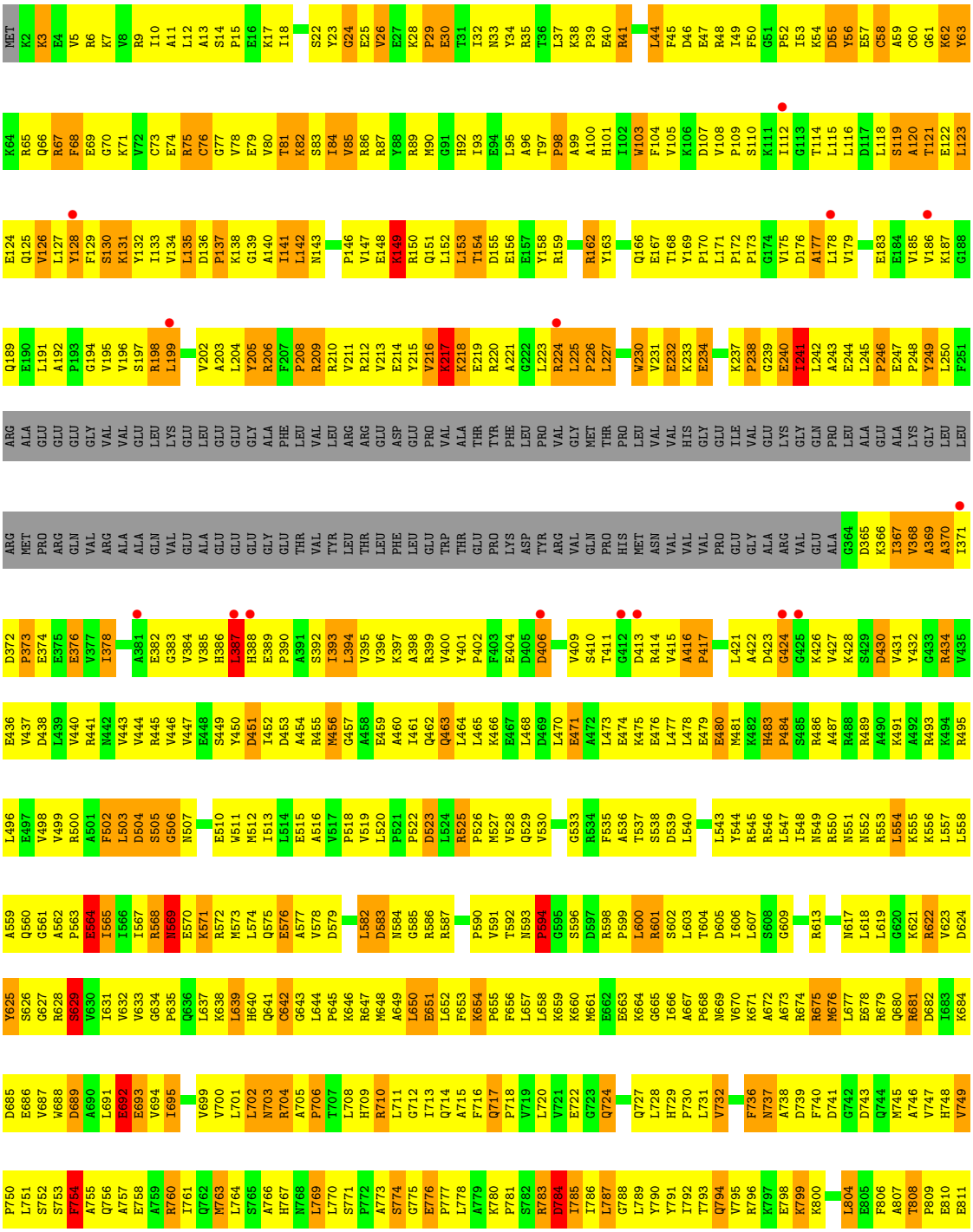


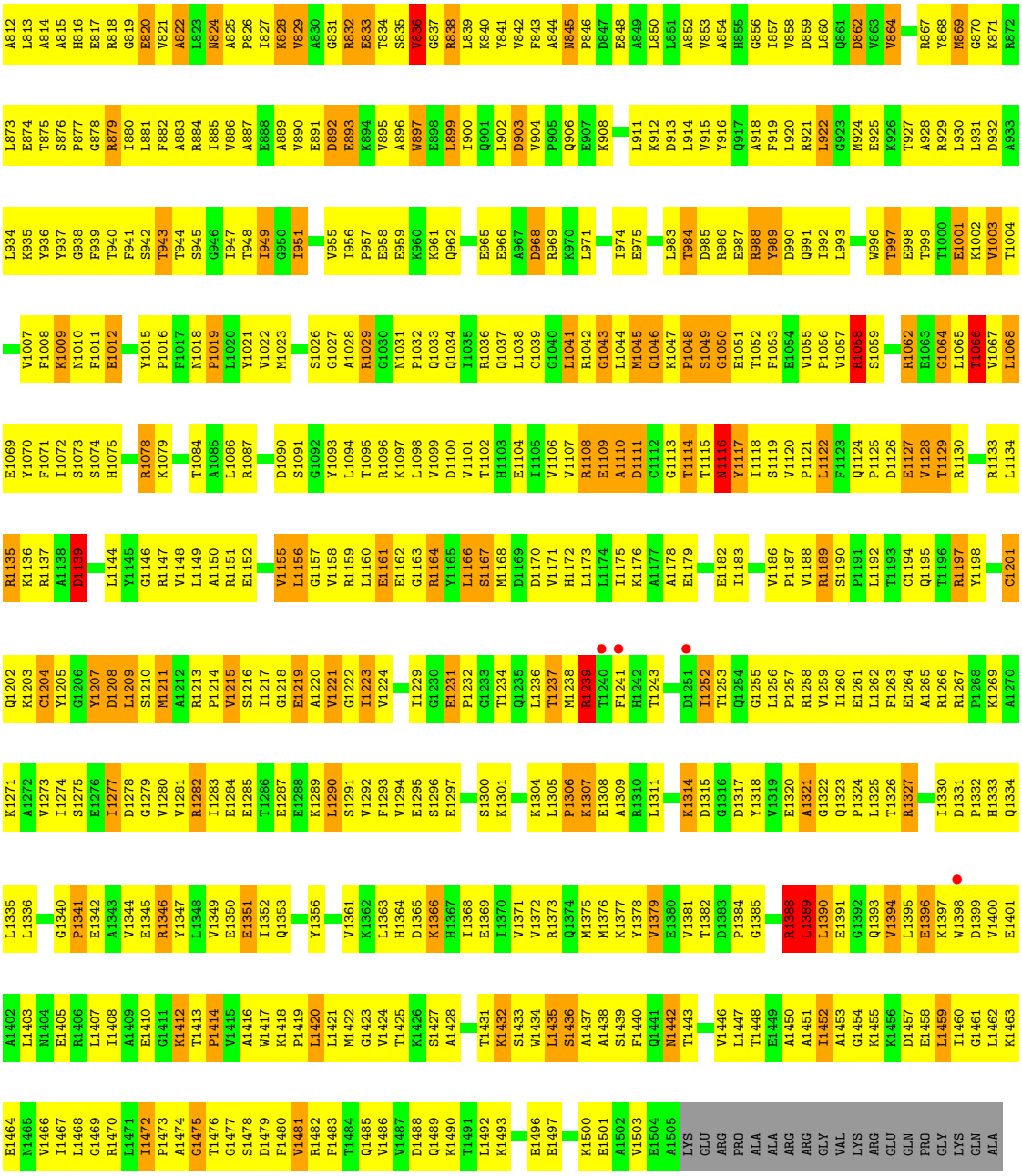




• Molecule 3: DNA-directed RNA polymerase subunit beta' chain

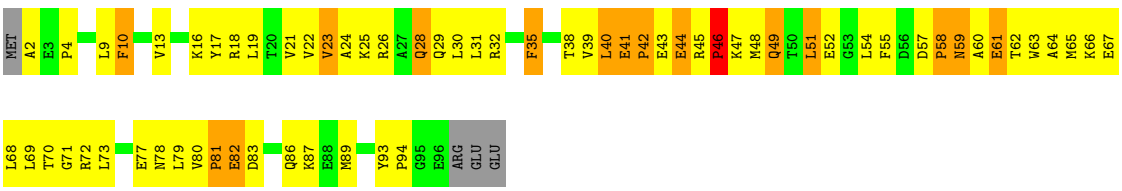






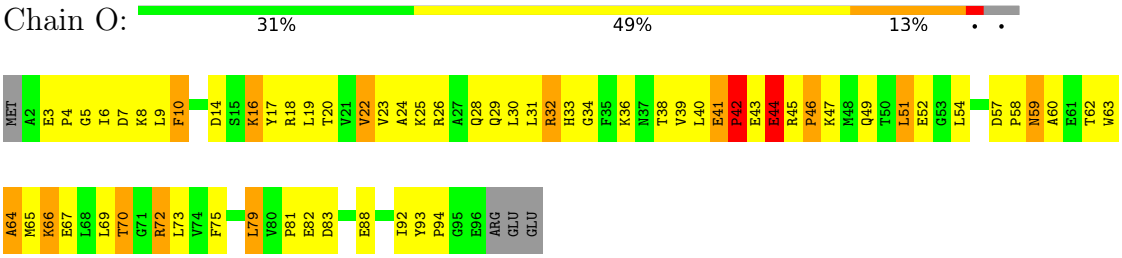
• Molecule 4: DNA-directed RNA polymerase omega chain

Chain E:



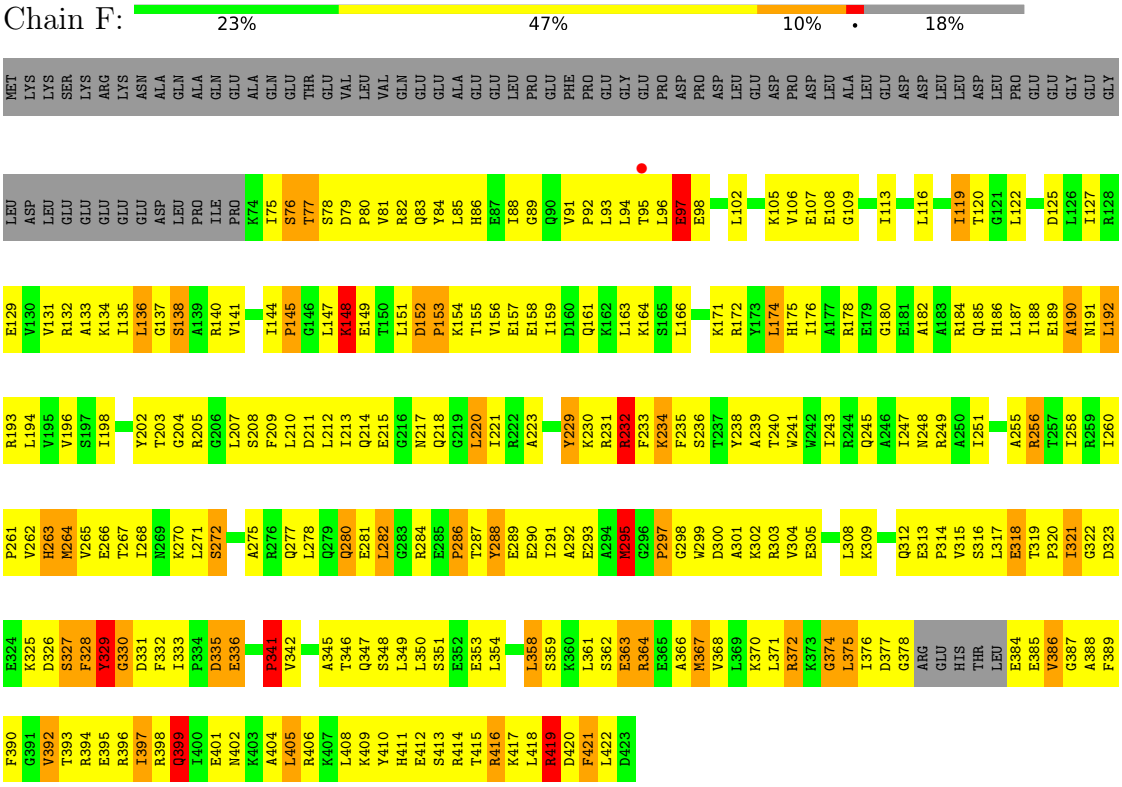
• Molecule 4: DNA-directed RNA polymerase omega chain

Chain O:



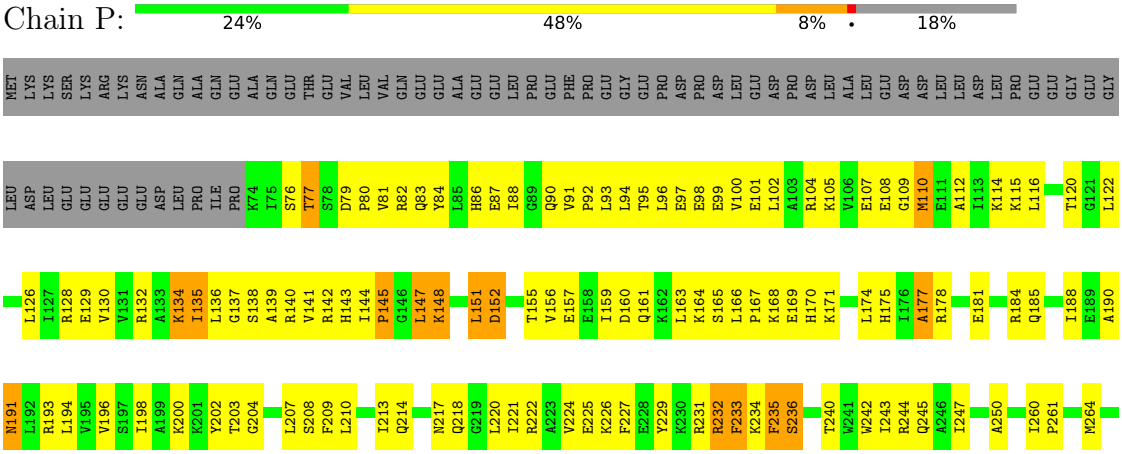
• Molecule 5: DNA-directed RNA polymerase sigma chain

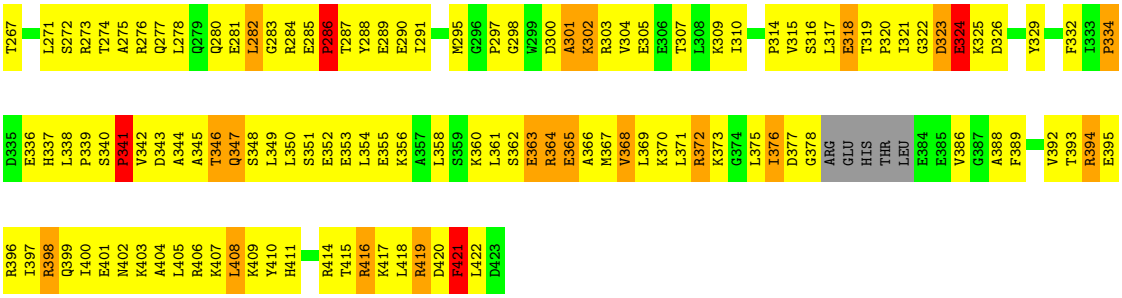
Chain F:



• Molecule 5: DNA-directed RNA polymerase sigma chain

Chain P:




T267
L271 S272 R273 T274 A275 R276 Q277 L278 Q279 Q280 E281 L282 G283 R284 E285 P286 T287 Y288 E289 E290 I291
M295 G296 P297 G298 W299 D300 A301 K302 R303 V304 E305 E306 T307 L308 K309 I310
P314 V315 S316 L317 E318 T319 P320 I321 G322 D323 E324 K325 D326
Y329
F332 I333 P334
D335 E336 H337 L338 P339 S340 P341 V342 D343 A344 A345 T346 Q347 S348 L349 L350 S351 E352 E353 L354 E355 K356 A357 L358 S359 K360 L361 S362 E363 R364 E365 A366 M367 V368 L369 K370 L371 R372 K373 G374 L375 I376 D377 G378 ARG GLU HIS THR LEU E384 E385 V386 G387 A388 F389
V392 T393 R394 E395
R396 I397 R398 Q399 I400 E401 M402 K403 A404 L405 R406 K407 L408 K409 Y410 H411
R414 T415 R416 K417 L418 R419 D420 F421 L422 D423

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α, β, γ	237.00Å 237.00Å 250.00Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 3.00 29.89 – 3.00	Depositor EDS
% Data completeness (in resolution range)	77.9 (30.00-3.00) 42.8 (29.89-3.00)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.12 (at 3.00Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.261 , 0.281 0.258 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	71.1	Xtriage
Anisotropy	0.354	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.23 , -24.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	0.499 for -h,-k,l 0.499 for h,-h-k,-l 0.044 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	54048	wwPDB-VP
Average B, all atoms (Å^2)	71.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.81% 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: STD, 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	0.43	0/1838	0.76	0/2498
1	B	0.38	0/1838	0.64	0/2498
1	K	0.46	0/1838	0.75	0/2498
1	L	0.39	0/1838	0.66	0/2498
2	C	0.45	0/8997	0.76	5/12164 (0.0%)
2	M	0.44	0/8997	0.76	5/12164 (0.0%)
3	D	0.46	0/11165	0.78	13/15088 (0.1%)
3	N	0.45	0/11165	0.78	14/15088 (0.1%)
4	E	0.40	0/783	0.77	2/1054 (0.2%)
4	O	0.44	0/783	0.82	1/1054 (0.1%)
5	F	0.41	0/2836	0.70	1/3812 (0.0%)
5	P	0.41	0/2836	0.69	0/3812
All	All	0.44	0/54914	0.76	41/74228 (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
2	C	0	1
3	D	0	1
3	N	0	1
All	All	0	3

There are no bond length outliers.

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	D	81	THR	N-CA-C	-7.96	89.51	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	M	728	HIS	N-CA-C	7.59	131.49	111.00
4	O	49	GLN	N-CA-C	7.23	130.53	111.00
3	N	1209	LEU	N-CA-C	-7.22	91.50	111.00
3	D	1209	LEU	N-CA-C	-7.04	92.00	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	258	TYR	Sidechain
3	D	132	TYR	Sidechain
3	N	625	TYR	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	1806	0	1861	279	0
1	B	1806	0	1861	211	0
1	K	1806	0	1861	232	0
1	L	1806	0	1861	210	0
2	C	8829	0	8933	1148	0
2	M	8829	0	8933	1178	0
3	D	10975	0	11213	1653	0
3	N	10975	0	11212	1616	0
4	E	769	0	775	99	0
4	O	769	0	775	87	0
5	F	2793	0	2873	320	0
5	P	2793	0	2873	362	0
6	D	43	0	44	10	0
6	M	43	0	44	8	0
7	D	2	0	0	0	0
7	N	2	0	0	0	0
8	D	1	0	0	0	0
8	N	1	0	0	0	0
All	All	54048	0	55119	6892	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 63.

The worst 5 of 6892 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:C:673:LEU:HD23	2:C:867:VAL:HA	1.22	1.20
3:D:141:ILE:H	3:D:141:ILE:HD12	1.11	1.16
3:N:172:PRO:HB3	3:N:178:LEU:HD22	1.22	1.16
2:C:857:ASP:HB2	2:C:978:ARG:HG2	1.29	1.15
3:D:1304:LYS:H	3:D:1304:LYS:HD3	1.03	1.15

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	227/315 (72%)	163 (72%)	48 (21%)	16 (7%)	1	6
1	B	227/315 (72%)	167 (74%)	46 (20%)	14 (6%)	1	8
1	K	227/315 (72%)	154 (68%)	45 (20%)	28 (12%)	0	1
1	L	227/315 (72%)	169 (74%)	41 (18%)	17 (8%)	1	5
2	C	1117/1119 (100%)	768 (69%)	229 (20%)	120 (11%)	0	2
2	M	1117/1119 (100%)	758 (68%)	222 (20%)	137 (12%)	0	1
3	D	1388/1524 (91%)	940 (68%)	286 (21%)	162 (12%)	0	1
3	N	1388/1524 (91%)	916 (66%)	317 (23%)	155 (11%)	0	2
4	E	93/99 (94%)	66 (71%)	17 (18%)	10 (11%)	0	2
4	O	93/99 (94%)	56 (60%)	25 (27%)	12 (13%)	0	1
5	F	341/423 (81%)	239 (70%)	57 (17%)	45 (13%)	0	1
5	P	341/423 (81%)	250 (73%)	54 (16%)	37 (11%)	0	2
All	All	6786/7590 (89%)	4646 (68%)	1387 (20%)	753 (11%)	0	2

5 of 753 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	118	ALA
1	A	160	ASP
1	A	188	GLN
1	B	3	ASP
1	B	96	THR

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	202/273 (74%)	171 (85%)	31 (15%)	2	13
1	B	202/273 (74%)	178 (88%)	24 (12%)	5	22
1	K	202/273 (74%)	172 (85%)	30 (15%)	3	14
1	L	202/273 (74%)	175 (87%)	27 (13%)	4	17
2	C	941/941 (100%)	829 (88%)	112 (12%)	5	22
2	M	941/941 (100%)	816 (87%)	125 (13%)	4	17
3	D	1170/1279 (92%)	972 (83%)	198 (17%)	2	11
3	N	1170/1279 (92%)	1000 (86%)	170 (14%)	3	15
4	E	83/87 (95%)	71 (86%)	12 (14%)	3	15
4	O	83/87 (95%)	72 (87%)	11 (13%)	4	17
5	F	300/370 (81%)	261 (87%)	39 (13%)	4	19
5	P	300/370 (81%)	281 (94%)	19 (6%)	18	51
All	All	5796/6446 (90%)	4998 (86%)	798 (14%)	3	17

5 of 798 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	L	140	MET
2	M	881	ASN
2	M	24	GLU
1	L	126	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	M	420	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 168 such sidechains are listed below:

Mol	Chain	Res	Type
2	M	704	HIS
3	N	768	ASN
2	M	845	ASN
3	N	125	GLN
3	N	1184	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 8 ligands modelled in this entry, 6 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
6	STD	D	1525	-	42,47,47	2.12	13 (30%)	47,73,73	1.74	9 (19%)
6	STD	M	1120	-	42,47,47	1.93	12 (28%)	47,73,73	2.01	12 (25%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	STD	D	1525	-	-	7/31/101/101	0/6/5/5
6	STD	M	1120	-	-	12/31/101/101	0/6/5/5

The worst 5 of 25 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	D	1525	STD	C22-N2	4.98	1.40	1.33
6	D	1525	STD	C26-C25	4.63	1.60	1.52
6	D	1525	STD	O5-C19	4.46	1.47	1.43
6	D	1525	STD	O8-C19	4.16	1.46	1.43
6	M	1120	STD	C26-C25	4.12	1.59	1.52

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
6	M	1120	STD	O4-C4-N1	6.75	113.28	105.92
6	D	1525	STD	C29-C19-C28	-5.62	108.73	113.30
6	M	1120	STD	C29-C19-C28	-5.16	109.10	113.30
6	M	1120	STD	O8-C19-C29	5.00	109.86	105.64
6	D	1525	STD	C19-O5-C13	4.56	117.72	112.80

There are no chirality outliers.

5 of 19 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	D	1525	STD	N1-C20-C21-C22
6	D	1525	STD	N1-C20-C21-C23
6	D	1525	STD	C3-C20-C21-C23
6	M	1120	STD	O4-C4-N1-C20
6	M	1120	STD	C9-C10-C13-C16

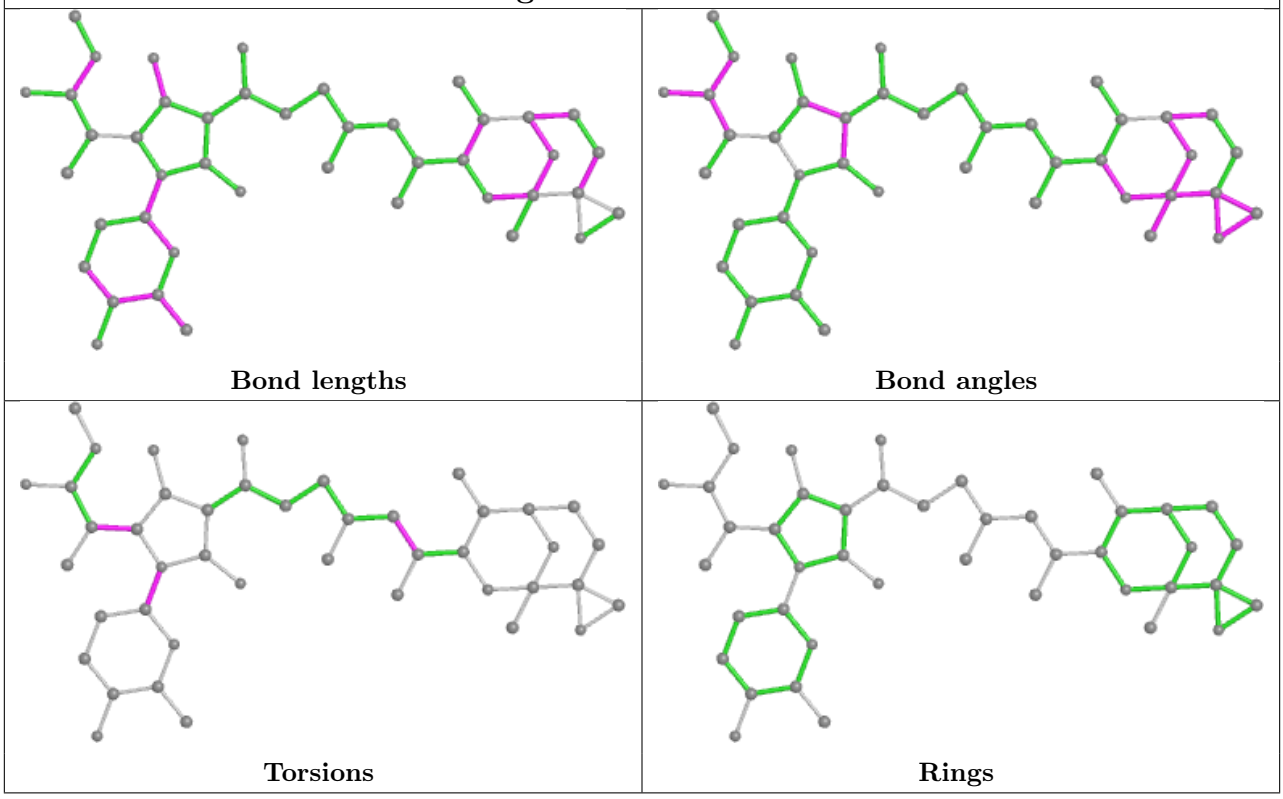
There are no ring outliers.

2 monomers are involved in 18 short contacts:

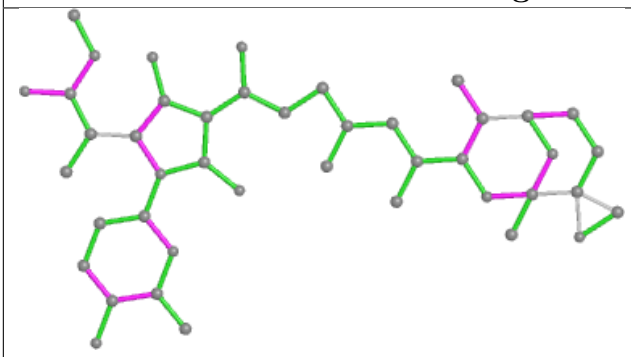
Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	D	1525	STD	10	0
6	M	1120	STD	8	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

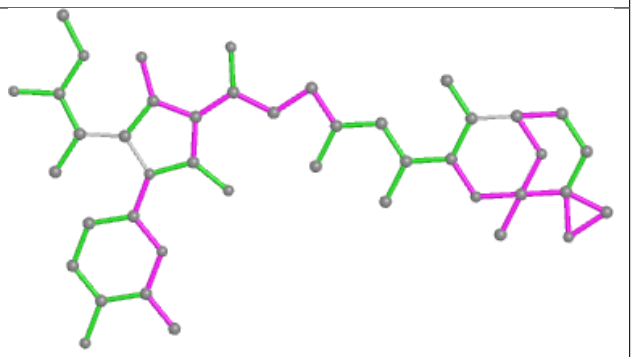
Ligand STD D 1525



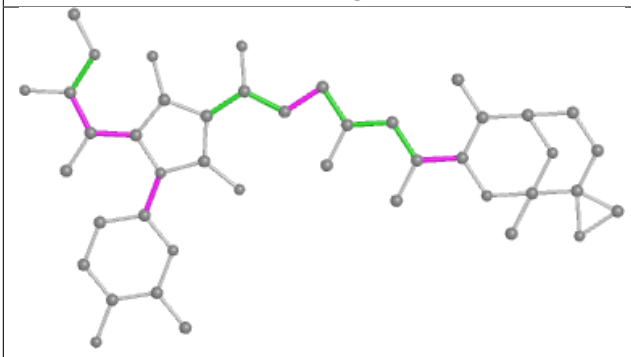
Ligand STD M 1120



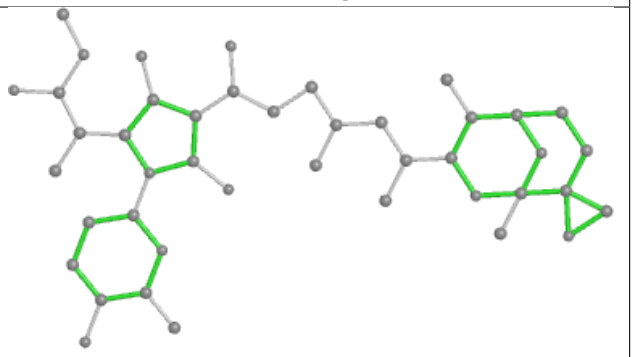
Bond lengths



Bond angles



Torsions



Rings

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	229/315 (72%)	-0.58	0 100 100	32, 58, 89, 110	0
1	B	229/315 (72%)	-0.52	1 (0%) 92 79	45, 96, 120, 124	0
1	K	229/315 (72%)	-0.56	0 100 100	24, 57, 82, 111	0
1	L	229/315 (72%)	-0.62	0 100 100	43, 78, 96, 117	0
2	C	1119/1119 (100%)	-0.55	5 (0%) 92 79	21, 67, 123, 137	0
2	M	1119/1119 (100%)	-0.54	3 (0%) 94 84	21, 69, 120, 130	0
3	D	1392/1524 (91%)	-0.53	10 (0%) 87 69	20, 64, 107, 125	0
3	N	1392/1524 (91%)	-0.48	19 (1%) 75 49	20, 64, 125, 140	0
4	E	95/99 (95%)	-0.56	0 100 100	52, 82, 99, 102	0
4	O	95/99 (95%)	-0.52	0 100 100	46, 86, 114, 119	0
5	F	345/423 (81%)	-0.57	1 (0%) 94 84	53, 77, 97, 104	0
5	P	345/423 (81%)	-0.57	0 100 100	47, 81, 97, 108	0
All	All	6818/7590 (89%)	-0.53	39 (0%) 89 72	20, 69, 118, 140	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	N	406	ASP	5.8
3	D	802	ALA	5.5
3	D	205	TYR	5.4
3	D	801	GLY	4.9
3	N	224	ARG	4.3

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

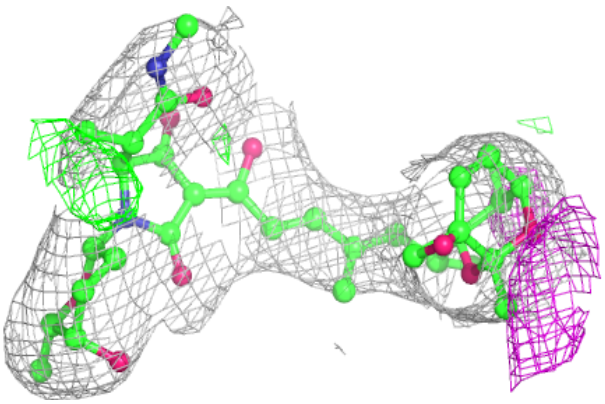
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

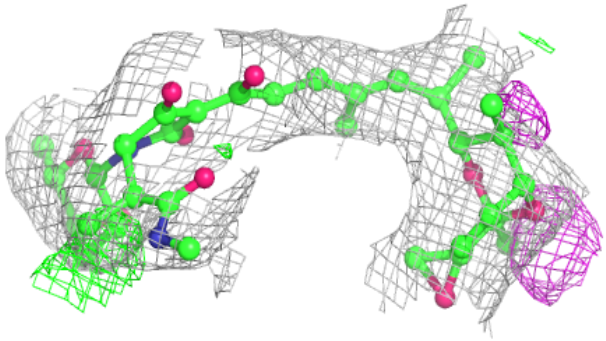
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	STD	M	1120	43/43	0.92	0.20	39,51,53,56	0
6	STD	D	1525	43/43	0.95	0.17	39,59,62,63	0
8	MG	N	9902	1/1	0.96	0.22	30,30,30,30	0
8	MG	D	9901	1/1	0.97	0.13	20,20,20,20	0
7	ZN	D	9002	1/1	0.98	0.22	47,47,47,47	0
7	ZN	N	9004	1/1	0.99	0.23	56,56,56,56	0
7	ZN	D	9001	1/1	0.99	0.20	56,56,56,56	0
7	ZN	N	9003	1/1	0.99	0.24	55,55,55,55	0

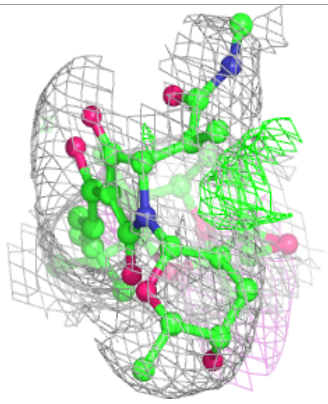
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around STD M 1120:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)

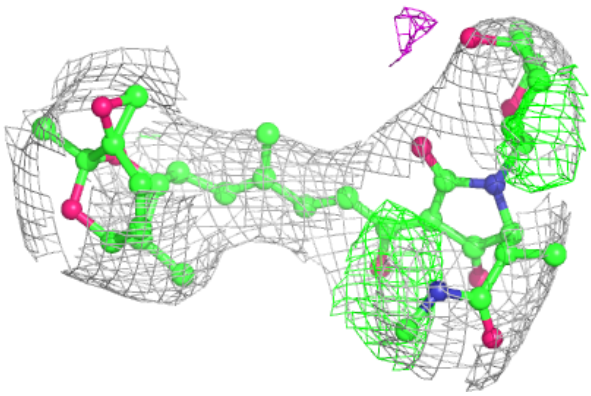


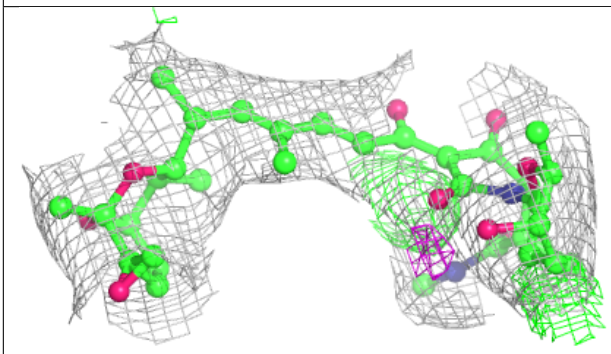


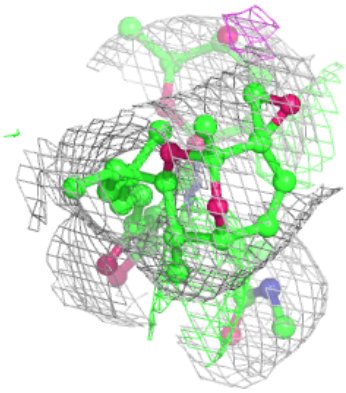


Electron density around STD D 1525:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)







6.5 Other polymers (i)

There are no such residues in this entry.