wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 09:17 AM GMT

PDB ID	:	3HKZ
Title	:	The X-ray crystal structure of RNA polymerase from Archaea
Authors	:	Murakami, K.S.
Deposited on	:	2009-05-26
Resolution	:	3.40 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
http://wwpdb.org/validation/2016/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:

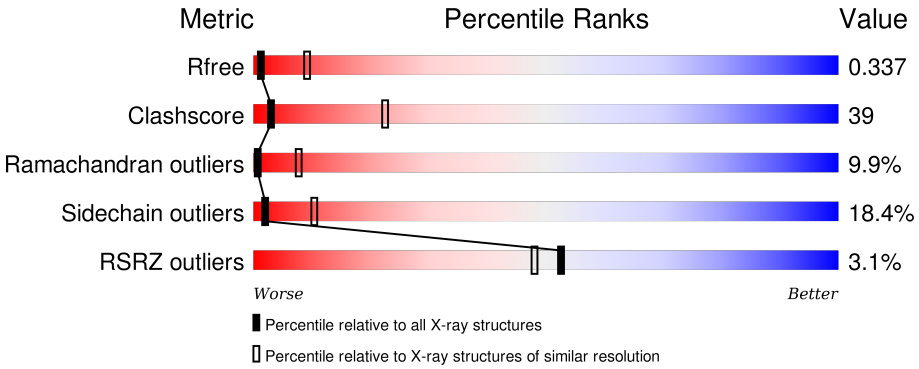
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 3.40 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}	91344	1476 (3.50-3.30)
Clashscore	102246	1611 (3.50-3.30)
Ramachandran outliers	100387	1571 (3.50-3.30)
Sidechain outliers	100360	1571 (3.50-3.30)
RSRZ outliers	91569	1485 (3.50-3.30)

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. 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	880	33% 46% 14% • 5%
1	I	880	2% 32% 48% 14% • 5%
2	C	395	6% 32% 43% 16% • 6%
2	M	395	4% 27% 44% 19% • 6%
3	B	1124	2% 32% 48% 15% • •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	J	1124	2% 31% 49% 15%
4	D	265	4% 54% 39% 6%
4	O	265	6% 52% 43%
5	E	180	4% 44% 41% 11%
5	Q	180	9% 42% 45% 10%
6	F	113	9% 41% 31% 5% 21%
6	R	113	4% 37% 36% 21%
7	G	132	3% 32% 42% 11% 14%
7	S	132	3% 28% 43% 14% 14%
8	H	84	29% 38% 17% 5% 12%
8	T	84	% 35% 33% 15% 5% 12%
9	K	95	% 27% 33% 21% 5% 14%
9	U	95	% 17% 44% 20% 5% 14%
10	L	92	% 41% 38% 20%
10	V	92	% 30% 54% 15%
11	N	66	3% 30% 55% 12%
11	W	66	2% 29% 55% 14%
12	P	48	2% 38% 42% 8% 10%
12	X	48	4% 38% 40% 10% 10%
13	Y	104	6% 23% 13% 7% 57%
13	Z	104	7% 23% 14% 6% 57%

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
14	ZN	B	2001	-	-	X	-
14	ZN	J	2001	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
16	F3S	O	1001	-	-	X	-

2 Entry composition

There are 16 unique types of molecules in this entry. The entry contains 53072 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 A'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	836	Total	C	N	O	S	0	0	0
			6673	4248	1178	1221	26			
1	I	836	Total	C	N	O	S	0	0	0
			6673	4248	1178	1221	26			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	370	Total	C	N	O	S	0	0	0
			2868	1818	490	551	9			
2	M	370	Total	C	N	O	S	0	0	0
			2868	1818	490	551	9			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	1	MET	-	EXPRESSION TAG	UNP P58192
C	2	GLU	-	EXPRESSION TAG	UNP P58192
C	3	GLY	-	EXPRESSION TAG	UNP P58192
M	1	MET	-	EXPRESSION TAG	UNP P58192
M	2	GLU	-	EXPRESSION TAG	UNP P58192
M	3	GLY	-	EXPRESSION TAG	UNP P58192

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	1090	Total	C	N	O	S	0	0	0
			8645	5483	1529	1602	31			
3	J	1090	Total	C	N	O	S	0	0	0
			8645	5483	1529	1602	31			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	264	Total	C	N	O	S	0	0	0
			2114	1355	342	403	14			
4	O	264	Total	C	N	O	S	0	0	0
			2114	1355	342	403	14			

- Molecule 5 is a protein called DNA-directed RNA polymerase, subunit E' (RpoE1).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	176	Total	C	N	O	S	0	0	0
			1402	903	236	259	4			
5	Q	176	Total	C	N	O	S	0	0	0
			1402	903	236	259	4			

- Molecule 6 is a protein called DNA-directed RNA polymerase, subunit F (RpoF).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	F	89	Total	C	N	O	S	0	0	0
			694	433	115	142	4			
6	R	89	Total	C	N	O	S	0	0	0
			694	433	115	142	4			

- Molecule 7 is a protein called DNA-directed RNA polymerase, subunit G (RpoG).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	G	113	Total	C	N	O	S	0	0	0
			884	556	149	174	5			
7	S	113	Total	C	N	O	S	0	0	0
			884	556	149	174	5			

- Molecule 8 is a protein called DNA-directed RNA polymerase subunit H.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
8	H	74	Total	C	N	O	0	0	0
			611	397	109	105			
8	T	74	Total	C	N	O	0	0	0
			611	397	109	105			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	K	82	Total	C	N	O	S	0	0	0
			658	420	121	116	1			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	U	82	Total	C	N	O	S	0	0	0
			658	420	121	116	1			

- Molecule 10 is a protein called DNA-directed RNA polymerase subunit L.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	L	92	Total	C	N	O	S	0	0	0
			723	459	121	141	2			
10	V	92	Total	C	N	O	S	0	0	0
			723	459	121	141	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	N	64	Total	C	N	O	S	0	0	0
			514	326	94	88	6			
11	W	64	Total	C	N	O	S	0	0	0
			514	326	94	88	6			

- Molecule 12 is a protein called DNA-directed RNA polymerase subunit P.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	P	43	Total	C	N	O	S	0	0	0
			346	230	58	53	5			
12	X	43	Total	C	N	O	S	0	0	0
			346	230	58	53	5			

- Molecule 13 is a protein called DNA-directed RNA polymerase subunit 13.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	Y	45	Total	C	N	O	S	0	0	0
			391	245	68	77	1			
13	Z	45	Total	C	N	O	S	0	0	0
			391	245	68	77	1			

- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	P	1	Total	Zn	0	0
			1	1		
14	J	1	Total	Zn	0	0
			1	1		

Continued on next page...

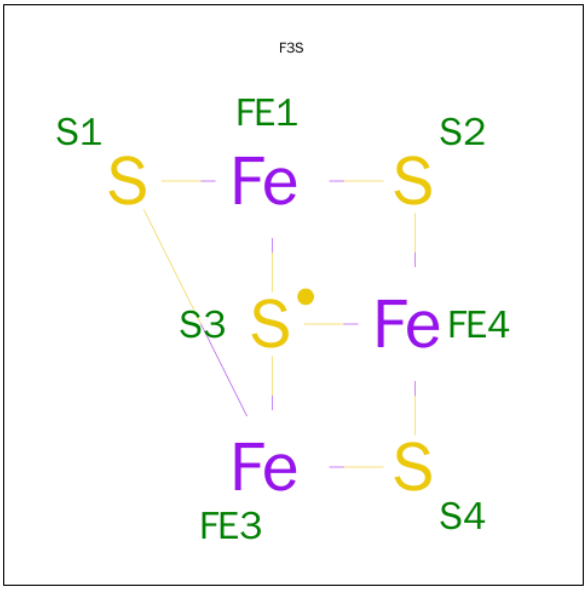
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
14	B	1	Total Zn 1 1	0	0
14	I	2	Total Zn 2 2	0	0
14	W	1	Total Zn 1 1	0	0
14	A	2	Total Zn 2 2	0	0
14	N	1	Total Zn 1 1	0	0
14	X	1	Total Zn 1 1	0	0

- Molecule 15 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
15	I	1	Total Mg 1 1	0	0
15	A	1	Total Mg 1 1	0	0

- Molecule 16 is FE3-S4 CLUSTER (three-letter code: F3S) (formula: Fe_3S_4).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
16	D	1	Total Fe S 7 3 4	0	0

Continued on next page...

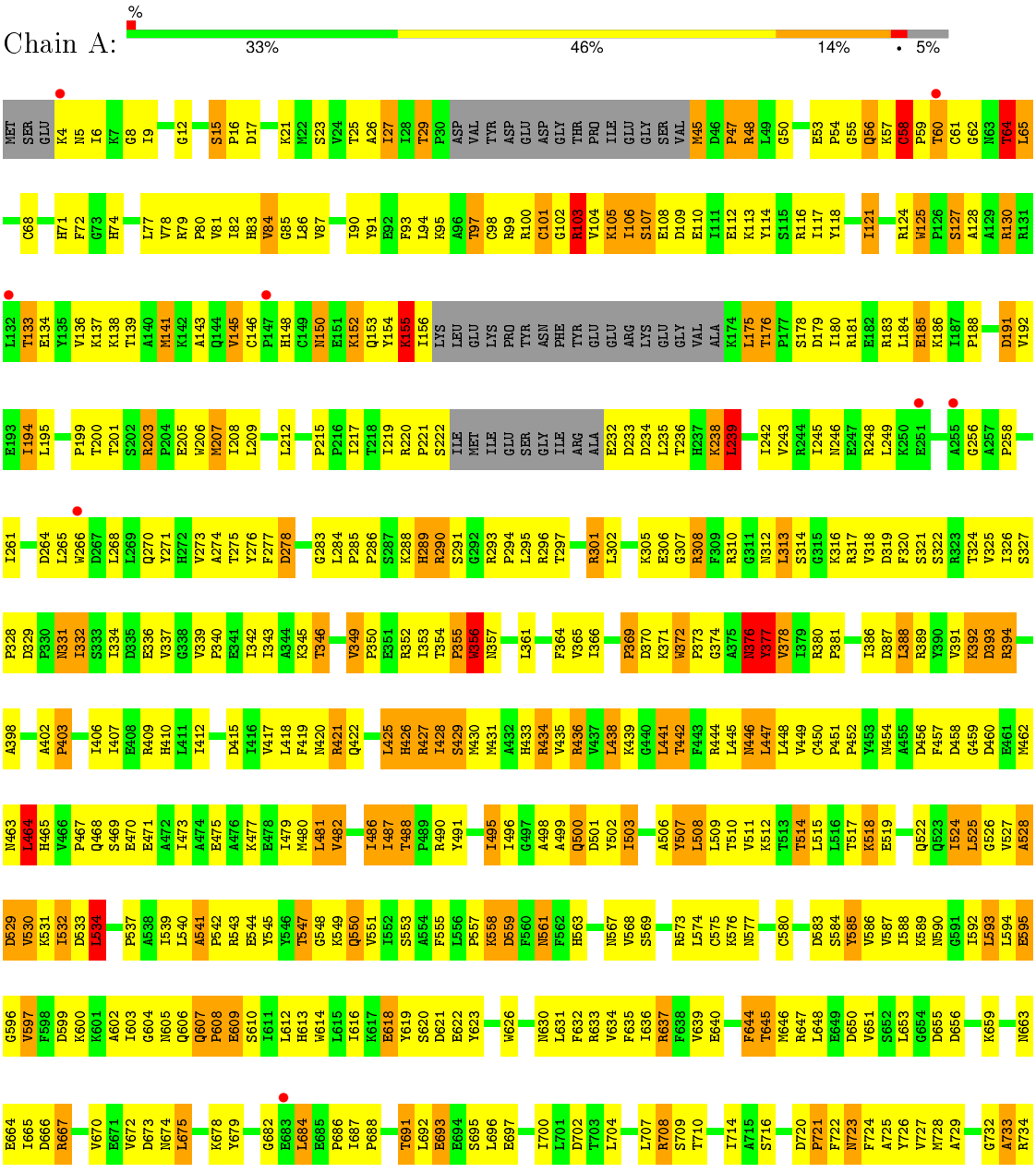
Continued from previous page...

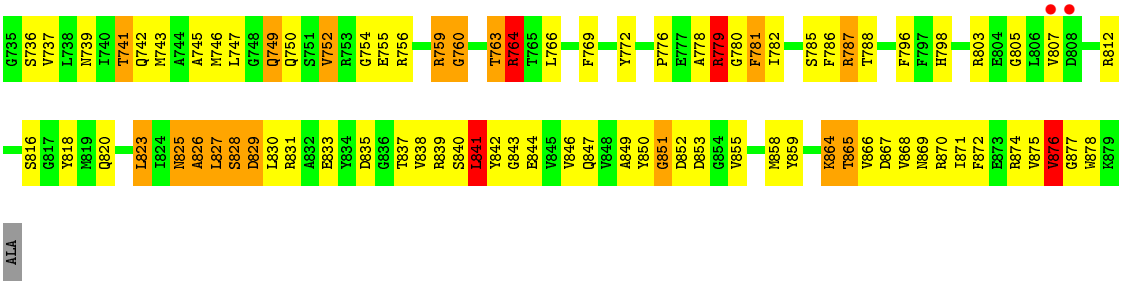
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	Fe	S		
16	O	1	7	3	4	0	0

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 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.

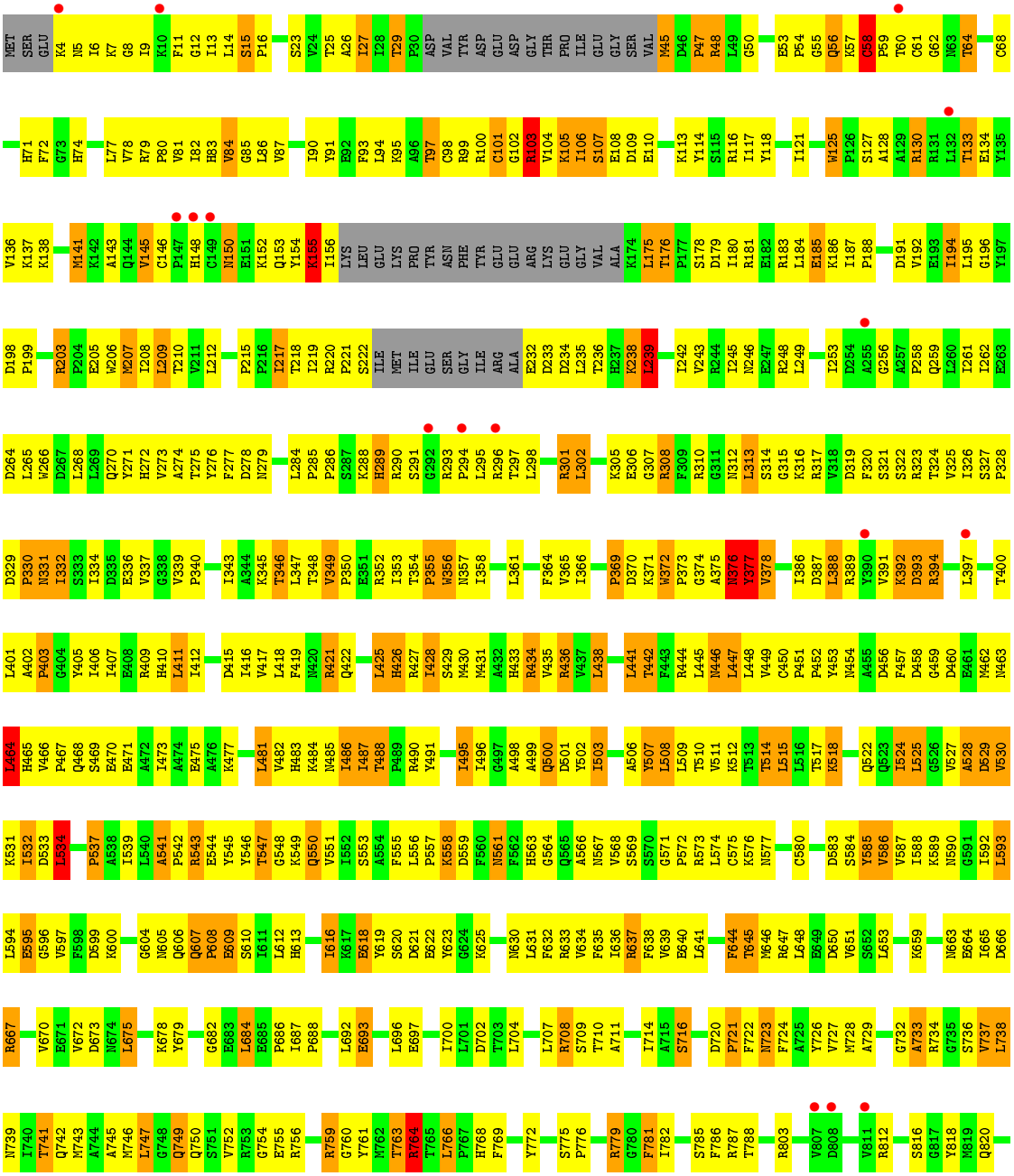
- Molecule 1: DNA-directed RNA polymerase subunit A'





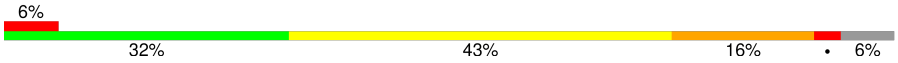
• Molecule 1: DNA-directed RNA polymerase subunit A'

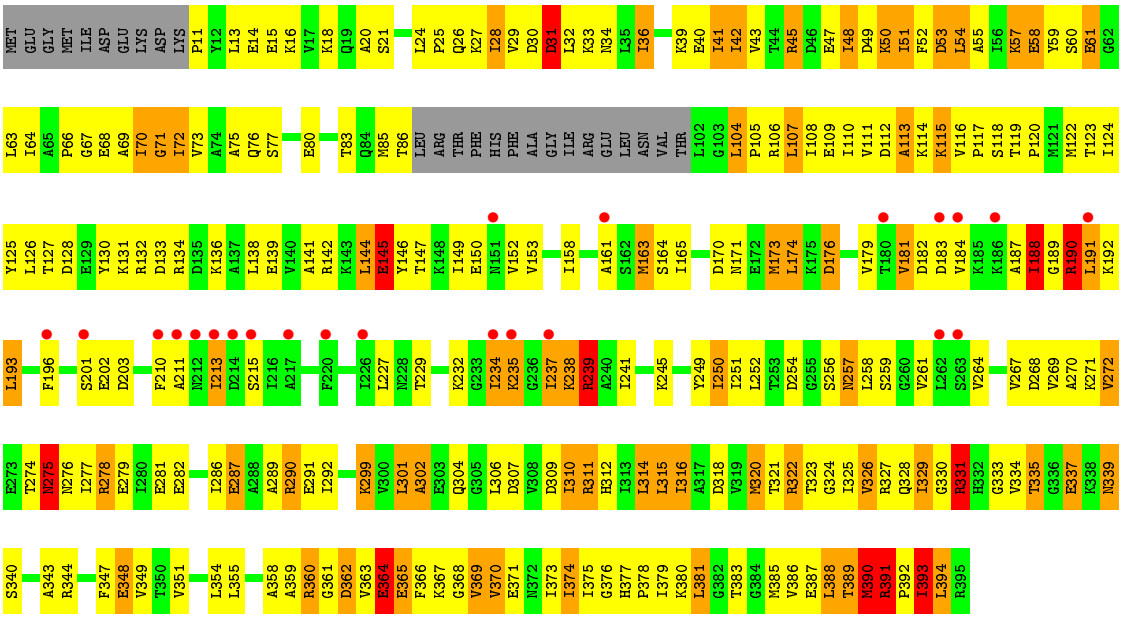
Chain I:



• Molecule 2: DNA-directed RNA polymerase subunit A"

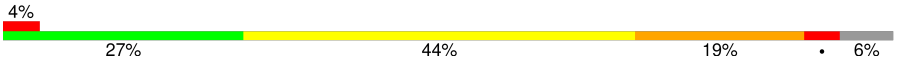
Chain C:

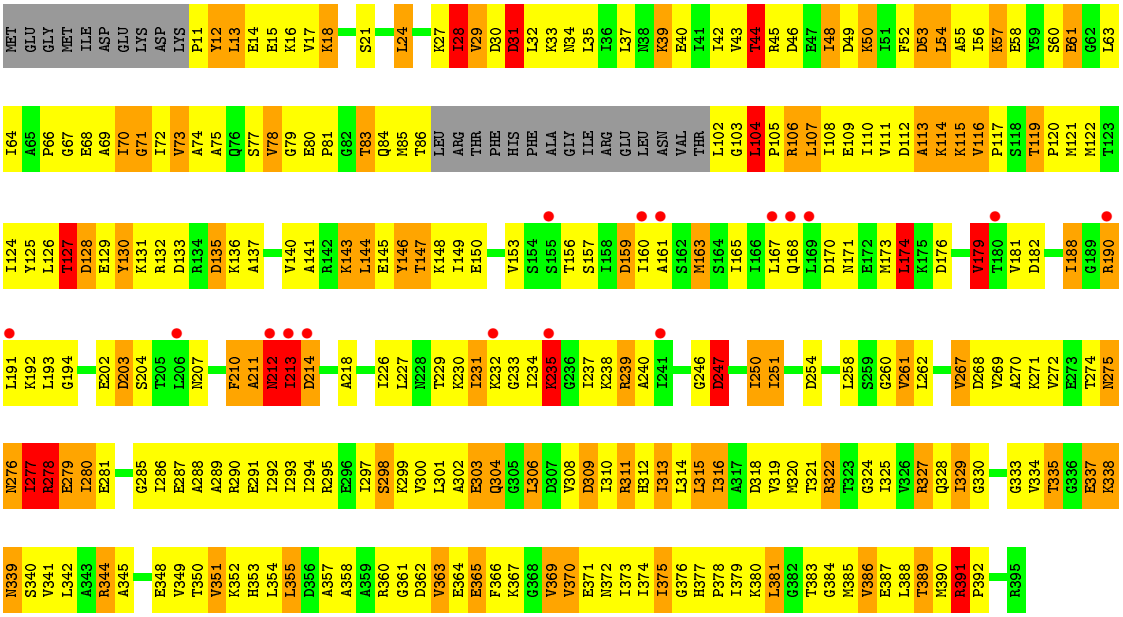




• Molecule 2: DNA-directed RNA polymerase subunit A"

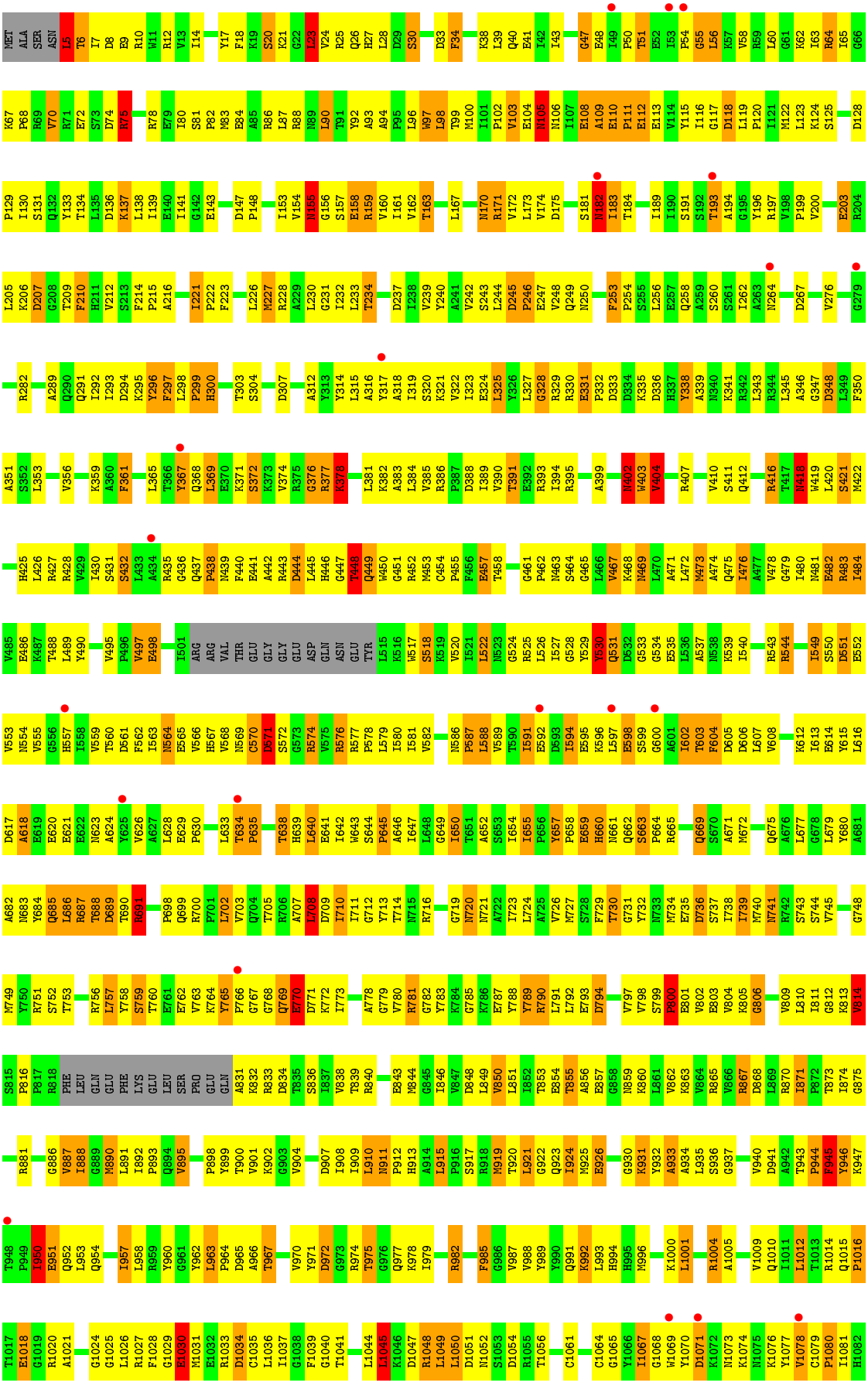
Chain M:





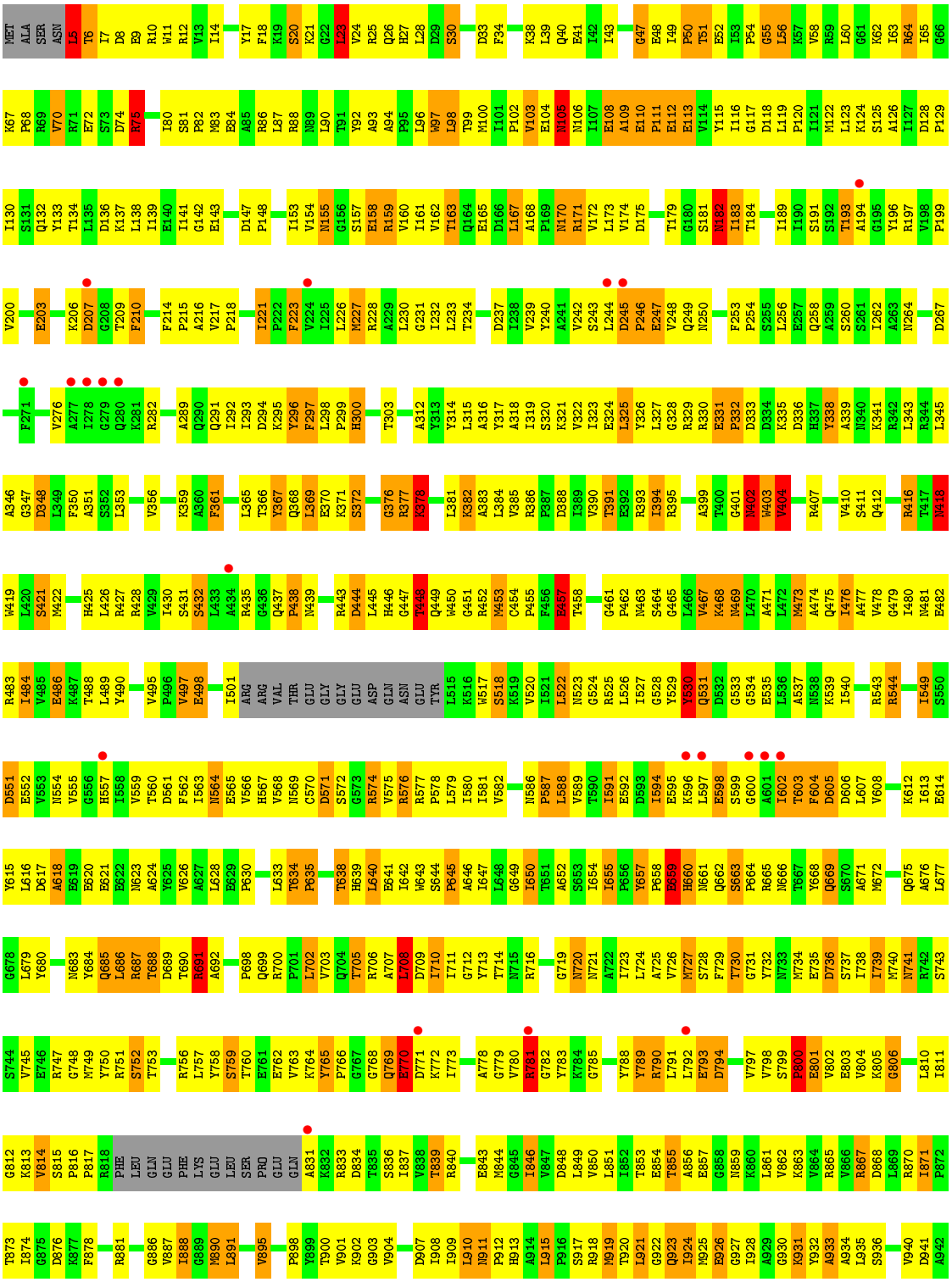
• Molecule 3: DNA-directed RNA polymerase subunit B

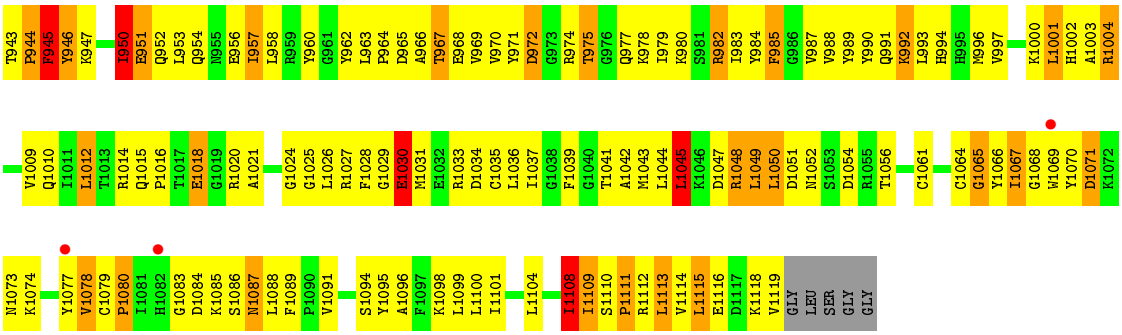
Chain B:


MET ALA SER ASN L5 T6 I7 D8 E9 R10 W11 R12 V13 I14 Y17 F18 K19 S20 K21 G22 L23 V24 R25 Q26 H27 L28 D29 S30 D33 F34 K38 L39 Q40 E41 I42 I43 G47 E48 I49 P50 T51 E52 I53 P54 G55 L56 K57 V58 R59 L60 G61 K62 I63 R64 I65 G66
K67 P68 R69 V70 R71 E72 S73 D74 R75 R78 E79 I80 S81 P82 M83 E84 A85 R86 I87 R88 N89 L90 T91 Y92 A93 A94 P95 L96 W97 L98 T99 M100 I101 P102 V103 E104 N105 N106 I107 E108 A109 E110 P111 E112 E113 V114 Y115 I116 G117 D118 L119 P120 I121 M122 L123 K124 S125 D128
P129 I130 S131 Q132 Y133 T134 L135 D136 K137 L138 I139 E140 I141 G142 E143 D147 P148 I153 V154 N155 G156 S157 E158 R159 V160 I161 V162 T163 L167 N170 R171 V172 L173 V174 D175 S181 N182 I183 T184 I189 I190 S191 S192 T193 A194 G195 Y196 R197 V198 P199 V200 E203 R204
L205 K206 D207 G208 T209 F210 H211 V212 S213 F214 P215 A216 I221 P222 F223 L226 M227 R228 A229 L230 G231 I232 L233 T234 D237 I238 V239 Y240 A241 V242 S243 L244 D245 P246 E247 V248 Q249 N250 F253 P254 S255 L256 E257 Q258 A259 S260 S261 I262 A263 N264 D267 V276 G279
R282 A289 Q290 Q291 I292 I293 D294 K295 Y296 F297 L298 P299 H300 T303 S304 D307 A312 Y313 Y314 L315 A316 Y317 A318 I319 S320 K321 V322 I323 E324 L325 Y326 L327 G328 R329 R330 E331 P332 D333 D334 K335 D336 H337 Y338 A339 N340 K341 R342 L343 R344 L345 A346 G347 D348 L349 F350
A351 S352 L353 V356 K359 A360 F361 L365 T366 Y367 Q368 L369 E370 K371 S372 K373 V374 R375 G376 R377 K378 L381 K382 A383 L384 V385 R386 P387 D388 I389 V390 T391 E392 R393 I394 R395 A399 N402 W403 V404 R407 V410 S411 Q412 R416 T417 N418 W419 L420 S421 M422
H425 L426 R427 R428 V429 I430 S431 S432 L433 A434 R435 G436 Q437 P438 N439 F440 E441 A442 R443 D444 L445 H446 G447 T448 Q449 W450 G451 R452 M453 C454 P455 F456 E457 T458 G461 P462 N463 S464 G465 L466 V467 K468 N469 L470 A471 L472 M473 A474 Q475 I476 A477 V478 G479 I480 N481 E482 R483 I484
V485 E486 K487 T488 L489 Y490 V495 P496 V497 E498 I501 ARG ARG VAL THR GLU GLY GLY GLU ASP GLN ASN GLU TYR L515 K516 W517 S518 K519 V520 I521 L522 N523 G524 R525 L526 I527 G528 Y529 Y530 Q531 D532 G533 G534 E535 L536 A537 N538 K539 I540 R543 R544 I549 S550 D551 E552
V553 N554 V555 G556 H557 I558 V559 T560 D561 F562 I563 N564 E565 V566 H567 V568 N569 C570 D571 S572 G573 R574 V575 R576 R577 P578 L579 I580 I581 V582 N586 P587 L588 V589 T590 I591 E592 D593 I594 E595 K596 L597 E598 S599 G600 A601 I602 T603 F604 D605 D606 L607 V608 K612 I613 E614 Y615 L616
D617 A618 E619 E620 E621 E622 N623 A624 Y625 V626 A627 L628 E629 P630 L633 T634 P635 T638 H639 L640 E641 I642 W643 S644 P645 A646 I647 L648 G649 I650 T651 A652 S653 I654 I655 P656 Y657 P658 E659 H660 M661 Q662 S663 P664 R665 A601 I602 T603 F604 D605 D606 L607 V608 Q669 S670 A671 H672 Q675 A676 L677 G678 L679 Y680 A681
A682 N683 Y684 Q685 L686 R687 T688 D689 T690 R691 P698 Q699 R700 F701 L702 V703 Q704 T705 R706 A707 L708 D709 I710 I711 G712 Y713 T714 N715 R716 G719 N720 N721 A722 I723 L724 A725 V726 M727 S728 F729 T730 G731 Y732 N733 M734 E735 D736 S737 I738 I739 M740 N741 R742 S743 S744 V745 G748
M749 Y750 R751 S752 T753 R756 L757 Y758 S759 T760 E761 E762 V763 K764 Y765 P766 G767 G768 Q769 E770 D771 K772 I773 A778 G779 V780 R781 G782 Y783 K784 G785 K786 E787 Y788 Y789 R790 L791 L792 E793 D794 V797 V798 S799 P800 E801 V802 E803 V804 K805 G806 V809 L810 I811 G812 K813 V814
S815 P816 P817 R818 PHE LEU GLN GLU PHE LYS GLU LEU SER PRO GLU GLN A831 K832 R833 D834 T835 S836 I837 V838 T839 R840 E843 M844 G845 I846 V847 D848 L849 V850 L851 I852 T853 E854 T855 A856 E857 G858 N859 K860 I861 V862 K863 V864 R865 V866 R867 D868 L869 R870 I871 P872 T873 I874 G875
R881 G886 V887 I888 G889 M890 L891 I892 P893 Q894 V895 P898 Y899 T900 V901 K902 G903 V904 D907 I908 I909 L910 N911 P912 H913 A914 L915 P916 S917 R918 M919 T920 L921 G922 Q923 I924 M925 E926 G930 K931 Y932 A933 A934 L935 S936 G937 V940 D941 A942 T943 P944 F945 Y946 K947
T948 P949 I950 E951 Q952 L953 Q954 I957 L958 R959 Y960 G961 Y962 L963 P964 D965 A966 T967 V970 Y971 D972 G973 R974 T975 G976 Q977 K978 I979 R982 F985 G986 V987 V988 Y989 Y990 Q991 K992 L993 H994 H995 M996 K1000 L1001 R1004 A1005 V1009 Q1010 I1011 L1012 T1013 R1014 Q1015 P1016
T1017 E1018 G1019 R1020 A1021 G1024 G1025 L1026 R1027 F1028 G1029 E1030 M1031 E1032 R1033 D1034 C1035 L1036 I1037 G1038 F1039 G1040 T1041 L1044 L1045 K1046 D1047 R1048 L1049 L1050 D1051 N1052 S1053 D1054 R1055 T1056 C1061 C1064 G1065 Y1066 I1067 G1068 W1069 Y1070 D1071 K1072 N1073 K1074 N1075 K1076 Y1077 V1078 C1079 P1080 I1081 H1082

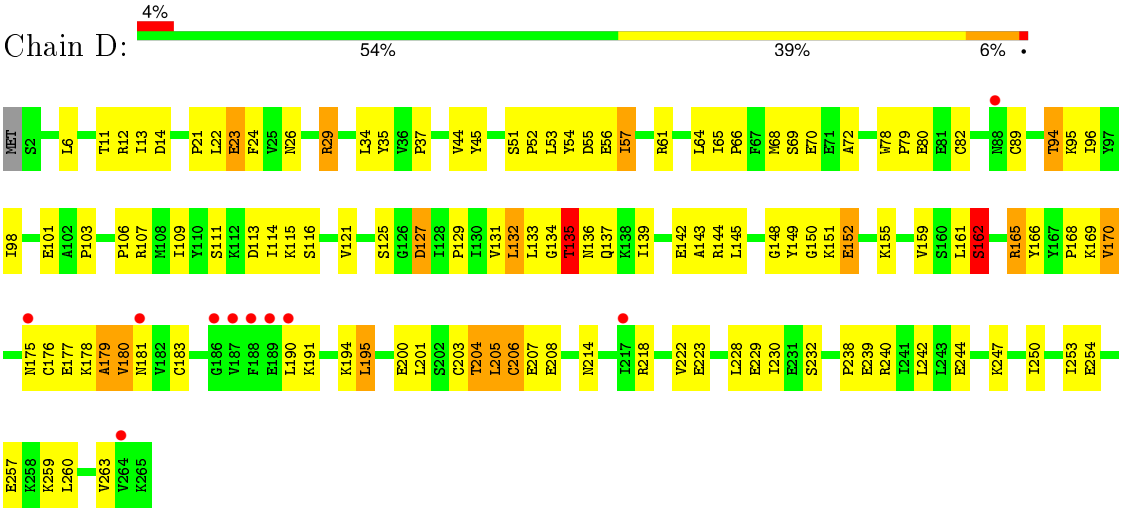
• Molecule 3: DNA-directed RNA polymerase subunit B

Chain J:

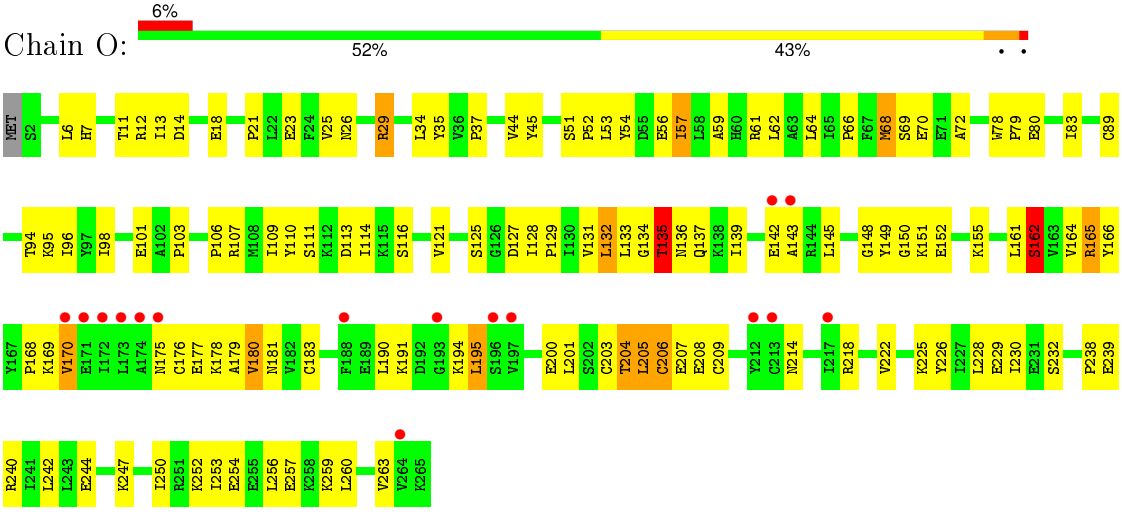




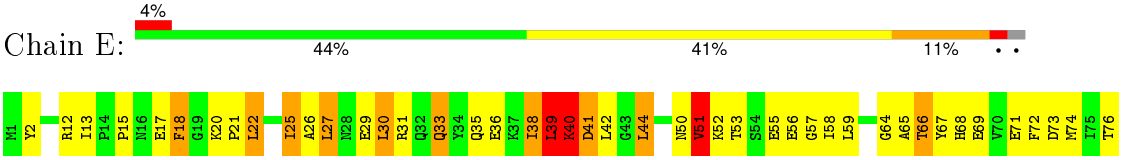
• Molecule 4: DNA-directed RNA polymerase subunit D

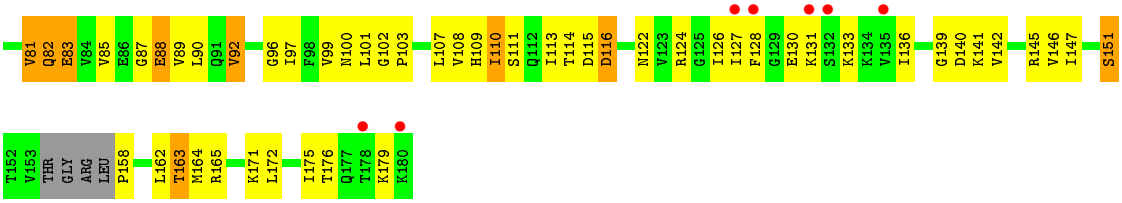


• Molecule 4: DNA-directed RNA polymerase subunit D

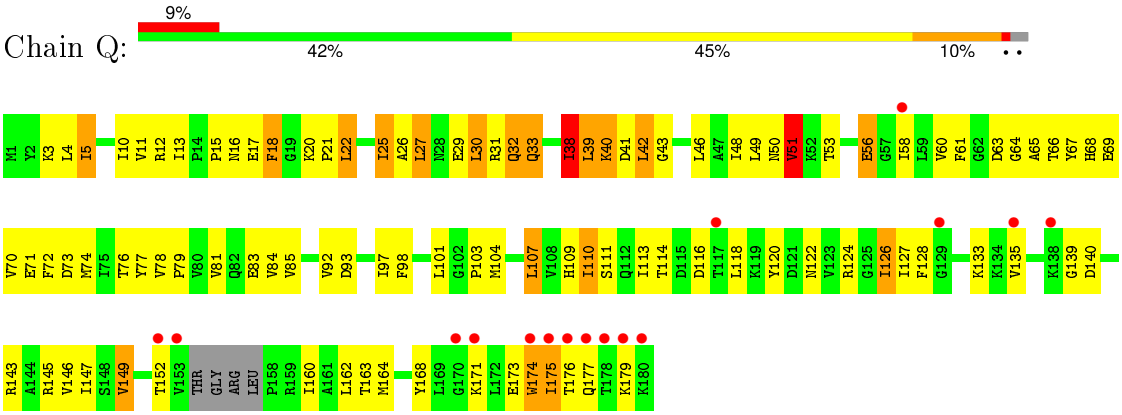


• Molecule 5: DNA-directed RNA polymerase, subunit E' (RpoE1)

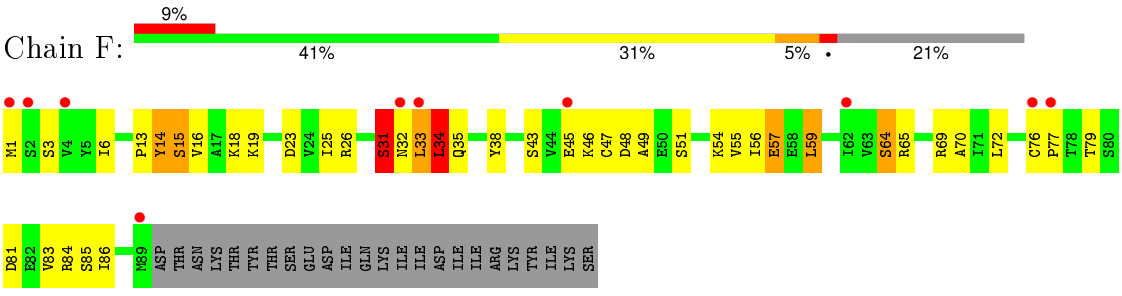




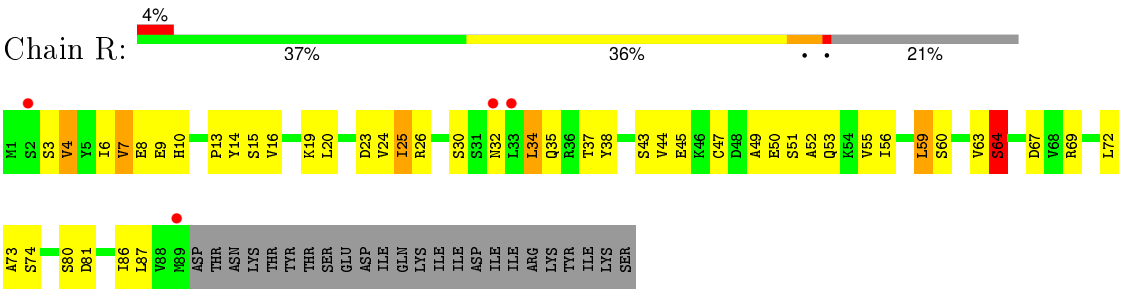
• Molecule 5: DNA-directed RNA polymerase, subunit E' (RpoE1)



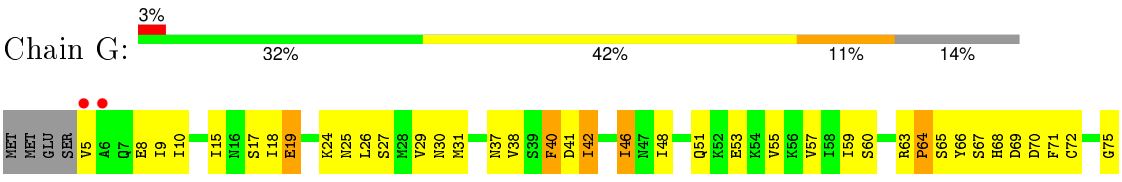
• Molecule 6: DNA-directed RNA polymerase, subunit F (RpoF)



• Molecule 6: DNA-directed RNA polymerase, subunit F (RpoF)

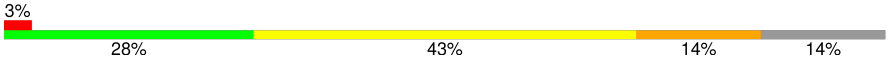


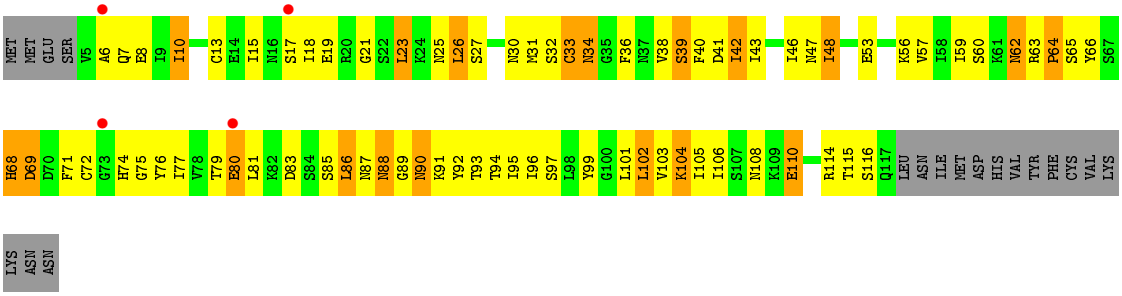
• Molecule 7: DNA-directed RNA polymerase, subunit G (RpoG)



• Molecule 7: DNA-directed RNA polymerase, subunit G (RpoG)

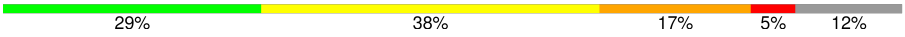
Chain S:

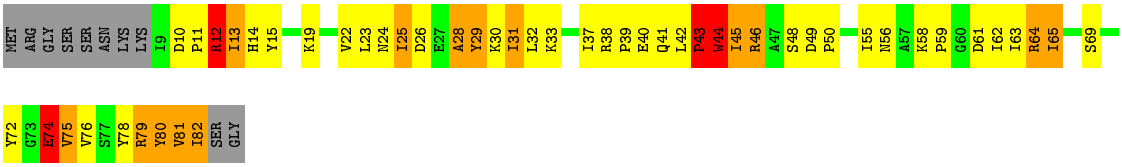




• Molecule 8: DNA-directed RNA polymerase subunit H

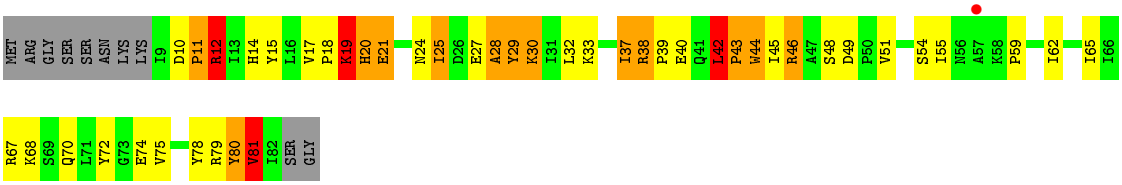
Chain H:





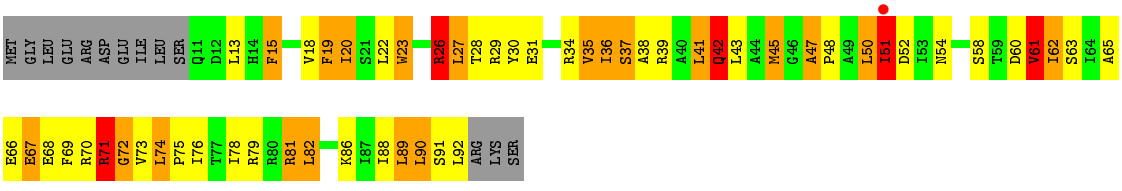
• Molecule 8: DNA-directed RNA polymerase subunit H

Chain T:



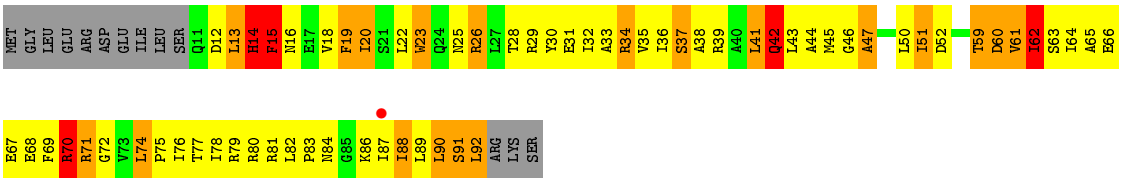
• Molecule 9: DNA-directed RNA polymerase subunit K

Chain K:



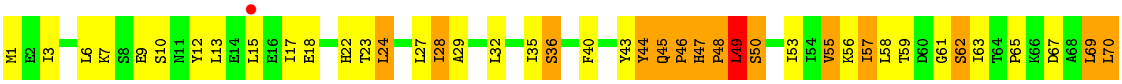
• Molecule 9: DNA-directed RNA polymerase subunit K

Chain U:



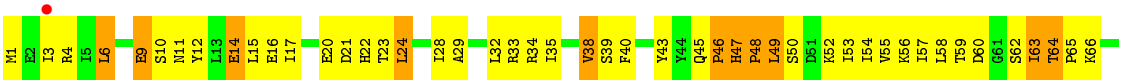
• Molecule 10: DNA-directed RNA polymerase subunit L

Chain L:



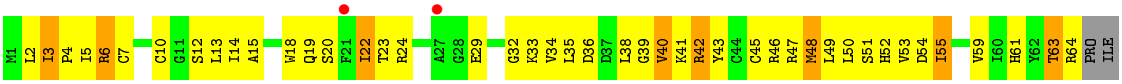
• Molecule 10: DNA-directed RNA polymerase subunit L

Chain V:



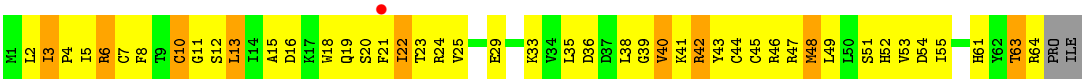
• Molecule 11: DNA-directed RNA polymerase subunit N

Chain N:



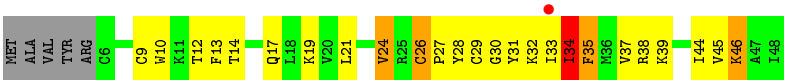
• Molecule 11: DNA-directed RNA polymerase subunit N

Chain W:



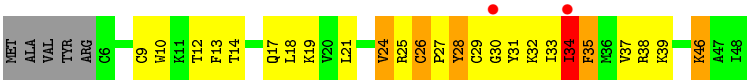
• Molecule 12: DNA-directed RNA polymerase subunit P

Chain P:

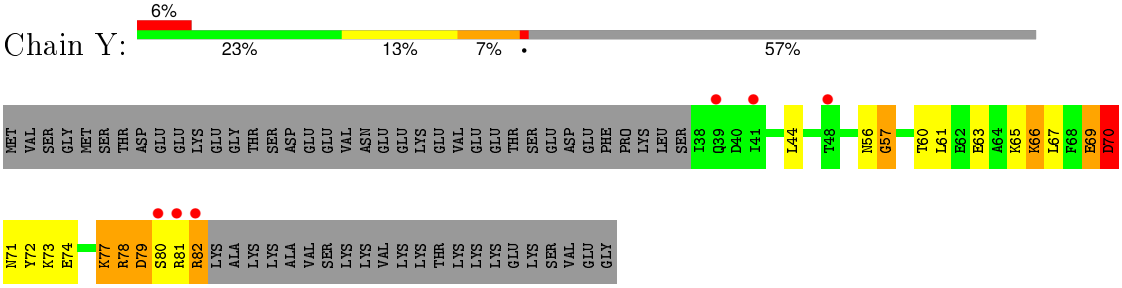


• Molecule 12: DNA-directed RNA polymerase subunit P

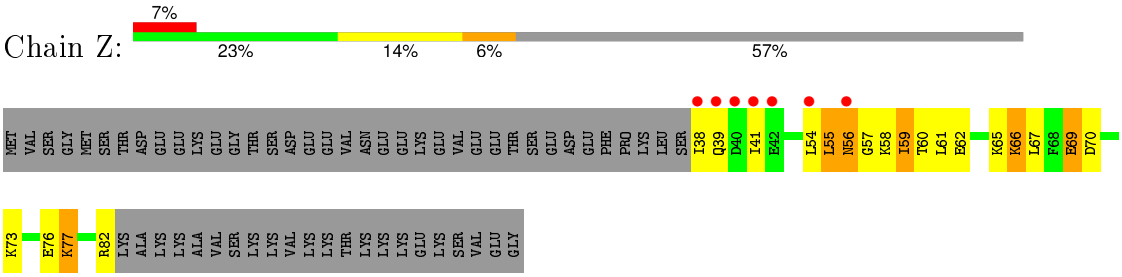
Chain X:



• Molecule 13: DNA-directed RNA polymerase subunit 13



• Molecule 13: DNA-directed RNA polymerase subunit 13



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	125.82Å 201.24Å 196.05Å 90.00° 100.92° 90.00°	Depositor
Resolution (Å)	40.00 – 3.40 39.79 – 3.40	Depositor EDS
% Data completeness (in resolution range)	80.3 (40.00-3.40) 80.3 (39.79-3.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.69 (at 3.40Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, R_{free}	0.265 , 0.341 0.264 , 0.337	Depositor DCC
R_{free} test set	5323 reflections (5.31%)	DCC
Wilson B-factor (Å^2)	76.2	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 71.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.41$, $< L^2 > = 0.24$	Xtriage
Outliers	0 of 105618 reflections	Xtriage
F_o,F_c correlation	0.85	EDS
Total number of atoms	53072	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.70% 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, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, F3S, 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	0.60	0/6815	0.76	4/9219 (0.0%)
1	I	0.62	0/6815	0.77	5/9219 (0.1%)
2	C	0.64	0/2892	0.85	1/3891 (0.0%)
2	M	0.66	0/2892	0.84	2/3891 (0.1%)
3	B	0.60	0/8810	0.78	7/11921 (0.1%)
3	J	0.60	0/8810	0.79	6/11921 (0.1%)
4	D	0.49	0/2152	0.66	0/2911
4	O	0.52	0/2152	0.66	0/2911
5	E	0.53	0/1423	0.72	0/1919
5	Q	0.52	0/1423	0.72	0/1919
6	F	0.50	0/701	0.66	1/949 (0.1%)
6	R	0.52	0/701	0.70	0/949
7	G	0.63	0/895	0.78	0/1203
7	S	0.68	0/895	0.75	1/1203 (0.1%)
8	H	0.54	0/625	0.77	0/848
8	T	0.61	0/625	0.83	1/848 (0.1%)
9	K	0.61	0/667	0.87	0/903
9	U	0.70	0/667	0.95	1/903 (0.1%)
10	L	0.51	0/733	0.71	1/986 (0.1%)
10	V	0.57	0/733	0.74	0/986
11	N	0.51	0/523	0.74	0/705
11	W	0.52	0/523	0.74	0/705
12	P	0.63	0/354	0.73	0/475
12	X	0.65	0/354	0.70	0/475
13	Y	0.71	0/395	0.73	0/527
13	Z	0.67	0/395	0.75	0/527
All	All	0.60	0/53970	0.77	30/72914 (0.0%)

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	M	0	1
3	B	0	1
3	J	0	1
7	G	0	1
9	K	0	1
9	U	0	1
All	All	0	6

There are no bond length outliers.

The worst 5 of 30 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	464	LEU	CA-CB-CG	6.85	131.04	115.30
1	I	239	LEU	CA-CB-CG	6.55	130.36	115.30
8	T	42	LEU	CA-CB-CG	6.31	129.80	115.30
1	A	239	LEU	CA-CB-CG	6.30	129.79	115.30
1	I	464	LEU	CA-CB-CG	6.19	129.55	115.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	B	945	PHE	Peptide
7	G	40	PHE	Peptide
3	J	945	PHE	Peptide
9	K	26	ARG	Peptide
2	M	27	LYS	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	6673	0	6763	612	0
1	I	6673	0	6763	619	0
2	C	2868	0	3035	277	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	M	2868	0	3035	331	0
3	B	8645	0	8780	817	0
3	J	8645	0	8780	862	0
4	D	2114	0	2146	120	0
4	O	2114	0	2146	122	0
5	E	1402	0	1467	80	0
5	Q	1402	0	1467	97	0
6	F	694	0	705	37	0
6	R	694	0	705	29	0
7	G	884	0	888	91	0
7	S	884	0	888	86	0
8	H	611	0	641	49	0
8	T	611	0	641	51	0
9	K	658	0	692	81	0
9	U	658	0	692	92	0
10	L	723	0	749	54	0
10	V	723	0	749	57	0
11	N	514	0	528	66	0
11	W	514	0	528	69	0
12	P	346	0	375	28	0
12	X	346	0	375	28	0
13	Y	391	0	389	27	0
13	Z	391	0	389	13	0
14	A	2	0	0	0	0
14	B	1	0	0	2	0
14	I	2	0	0	0	0
14	J	1	0	0	2	0
14	N	1	0	0	0	0
14	P	1	0	0	0	0
14	W	1	0	0	0	0
14	X	1	0	0	0	0
15	A	1	0	0	0	0
15	I	1	0	0	0	0
16	D	7	0	0	1	0
16	O	7	0	0	6	0
All	All	53072	0	54316	4172	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 39.

The worst 5 of 4172 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)
8:H:29:TYR:OH	13:Y:67:LEU:HD22	1.31	1.25
4:D:183:CYS:SG	16:D:1001:F3S:S2	2.42	1.17
6:R:72:LEU:HD21	6:R:86:ILE:HD12	1.30	1.14
5:Q:147:ILE:HD11	5:Q:163:THR:HB	1.21	1.14
4:D:250:ILE:HD11	10:L:84:ILE:HD11	1.31	1.13

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	828/880 (94%)	597 (72%)	154 (19%)	77 (9%)	1 9
1	I	828/880 (94%)	594 (72%)	157 (19%)	77 (9%)	1 9
2	C	366/395 (93%)	225 (62%)	92 (25%)	49 (13%)	0 4
2	M	366/395 (93%)	232 (63%)	84 (23%)	50 (14%)	0 3
3	B	1084/1124 (96%)	752 (69%)	215 (20%)	117 (11%)	0 6
3	J	1084/1124 (96%)	752 (69%)	217 (20%)	115 (11%)	0 6
4	D	262/265 (99%)	199 (76%)	50 (19%)	13 (5%)	3 24
4	O	262/265 (99%)	198 (76%)	53 (20%)	11 (4%)	3 29
5	E	172/180 (96%)	125 (73%)	35 (20%)	12 (7%)	1 15
5	Q	172/180 (96%)	128 (74%)	34 (20%)	10 (6%)	2 20
6	F	87/113 (77%)	57 (66%)	25 (29%)	5 (6%)	2 20
6	R	87/113 (77%)	59 (68%)	19 (22%)	9 (10%)	1 7
7	G	111/132 (84%)	82 (74%)	22 (20%)	7 (6%)	2 18
7	S	111/132 (84%)	84 (76%)	16 (14%)	11 (10%)	1 8
8	H	72/84 (86%)	46 (64%)	12 (17%)	14 (19%)	0 1
8	T	72/84 (86%)	39 (54%)	20 (28%)	13 (18%)	0 1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
9	K	80/95 (84%)	50 (62%)	21 (26%)	9 (11%)	0 6
9	U	80/95 (84%)	43 (54%)	18 (22%)	19 (24%)	0 1
10	L	90/92 (98%)	66 (73%)	16 (18%)	8 (9%)	1 9
10	V	90/92 (98%)	58 (64%)	28 (31%)	4 (4%)	3 28
11	N	62/66 (94%)	41 (66%)	17 (27%)	4 (6%)	1 17
11	W	62/66 (94%)	43 (69%)	15 (24%)	4 (6%)	1 17
12	P	41/48 (85%)	29 (71%)	8 (20%)	4 (10%)	1 8
12	X	41/48 (85%)	28 (68%)	8 (20%)	5 (12%)	0 5
13	Y	43/104 (41%)	26 (60%)	13 (30%)	4 (9%)	1 9
13	Z	43/104 (41%)	30 (70%)	11 (26%)	2 (5%)	3 26
All	All	6596/7156 (92%)	4583 (70%)	1360 (21%)	653 (10%)	1 8

5 of 653 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	47	PRO
1	A	56	GLN
1	A	58	CYS
1	A	64	THR
1	A	145	VAL

5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all 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	729/766 (95%)	601 (82%)	128 (18%)	2 12
1	I	729/766 (95%)	604 (83%)	125 (17%)	2 14
2	C	318/340 (94%)	241 (76%)	77 (24%)	1 4
2	M	318/340 (94%)	226 (71%)	92 (29%)	0 2
3	B	937/965 (97%)	772 (82%)	165 (18%)	2 12
3	J	937/965 (97%)	771 (82%)	166 (18%)	2 12

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	D	241/242 (100%)	220 (91%)	21 (9%)	13 47
4	O	241/242 (100%)	221 (92%)	20 (8%)	14 49
5	E	156/159 (98%)	128 (82%)	28 (18%)	2 11
5	Q	156/159 (98%)	123 (79%)	33 (21%)	1 6
6	F	82/106 (77%)	72 (88%)	10 (12%)	6 28
6	R	82/106 (77%)	72 (88%)	10 (12%)	6 28
7	G	105/125 (84%)	87 (83%)	18 (17%)	2 14
7	S	105/125 (84%)	82 (78%)	23 (22%)	1 6
8	H	67/75 (89%)	49 (73%)	18 (27%)	0 3
8	T	67/75 (89%)	54 (81%)	13 (19%)	2 8
9	K	72/84 (86%)	48 (67%)	24 (33%)	0 2
9	U	72/84 (86%)	50 (69%)	22 (31%)	0 2
10	L	81/81 (100%)	66 (82%)	15 (18%)	2 10
10	V	81/81 (100%)	63 (78%)	18 (22%)	1 5
11	N	58/60 (97%)	48 (83%)	10 (17%)	2 13
11	W	58/60 (97%)	48 (83%)	10 (17%)	2 13
12	P	39/43 (91%)	32 (82%)	7 (18%)	2 11
12	X	39/43 (91%)	32 (82%)	7 (18%)	2 11
13	Y	43/97 (44%)	34 (79%)	9 (21%)	1 6
13	Z	43/97 (44%)	34 (79%)	9 (21%)	1 6
All	All	5856/6286 (93%)	4778 (82%)	1078 (18%)	2 10

5 of 1078 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
9	K	89	LEU
1	I	488	THR
7	S	104	LYS
10	L	79	MET
1	I	130	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 134 such sidechains are listed below:

Mol	Chain	Res	Type
9	K	54	ASN
1	I	446	ASN
6	R	53	GLN
11	N	19	GLN
1	I	259	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 14 ligands modelled in this entry, 12 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	$\#\|Z\| > 2$
16	F3S	D	1001	4	0,9,9	0.00	-	0,15,15	0.00	-
16	F3S	O	1001	-	0,9,9	0.00	-	0,15,15	0.00	-

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
16	F3S	D	1001	4	-	0/0/24/24	0/0/3/3
16	F3S	O	1001	-	-	0/0/24/24	0/0/3/3

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.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	D	1001	F3S	1	0
16	O	1001	F3S	6	0

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	836/880 (95%)	-0.14	10 (1%) 81 75	28, 71, 105, 133	0
1	I	836/880 (95%)	-0.15	16 (1%) 70 64	28, 71, 106, 133	0
2	C	370/395 (93%)	0.21	23 (6%) 24 22	37, 78, 130, 138	0
2	M	370/395 (93%)	0.09	16 (4%) 39 34	38, 79, 130, 138	0
3	B	1090/1124 (96%)	-0.12	21 (1%) 70 64	33, 66, 117, 149	0
3	J	1090/1124 (96%)	-0.11	24 (2%) 65 60	33, 67, 117, 149	0
4	D	264/265 (99%)	0.11	10 (3%) 44 39	57, 86, 121, 139	0
4	O	264/265 (99%)	0.22	16 (6%) 25 23	57, 86, 121, 139	0
5	E	176/180 (97%)	0.20	7 (3%) 42 37	58, 98, 140, 143	0
5	Q	176/180 (97%)	0.26	16 (9%) 11 11	58, 98, 139, 142	0
6	F	89/113 (78%)	0.27	10 (11%) 7 7	100, 127, 132, 133	0
6	R	89/113 (78%)	0.24	4 (4%) 37 33	100, 127, 132, 133	0
7	G	113/132 (85%)	0.19	4 (3%) 48 42	65, 88, 108, 112	0
7	S	113/132 (85%)	0.12	4 (3%) 48 42	66, 88, 108, 112	0
8	H	74/84 (88%)	-0.11	0 100 100	66, 81, 99, 102	0
8	T	74/84 (88%)	-0.16	1 (1%) 78 73	66, 81, 99, 102	0
9	K	82/95 (86%)	-0.39	1 (1%) 81 75	41, 57, 86, 94	0
9	U	82/95 (86%)	-0.31	1 (1%) 81 75	43, 58, 87, 93	0
10	L	92/92 (100%)	-0.10	1 (1%) 82 77	53, 73, 103, 107	0
10	V	92/92 (100%)	-0.00	1 (1%) 82 77	54, 73, 103, 107	0
11	N	64/66 (96%)	-0.10	2 (3%) 52 48	62, 80, 103, 124	0
11	W	64/66 (96%)	-0.04	1 (1%) 74 69	62, 80, 103, 124	0
12	P	43/48 (89%)	0.10	1 (2%) 64 58	66, 89, 103, 110	0
12	X	43/48 (89%)	-0.08	2 (4%) 35 32	66, 89, 103, 110	0

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2		OWAB(Å^2)	Q<0.9
13	Y	45/104 (43%)	0.63	6 (13%)	4 4	98, 108, 124, 125	0
13	Z	45/104 (43%)	0.79	7 (15%)	3 3	98, 108, 124, 124	0
All	All	6676/7156 (93%)	-0.03	205 (3%)	52 48	28, 76, 127, 149	0

The worst 5 of 205 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	J	278	ILE	7.4
6	R	89	MET	5.5
2	C	213	ILE	5.2
3	B	434	ALA	5.0
13	Y	82	ARG	4.9

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
14	ZN	I	1001	1/1	0.99	0.12	-1.27	57,57,57,57	0
14	ZN	A	1002	1/1	0.98	0.06	-1.37	69,69,69,69	0
16	F3S	D	1001	7/7	0.99	0.10	-1.39	79,81,82,84	0
14	ZN	X	1001	1/1	0.98	0.09	-1.40	54,54,54,54	0
16	F3S	O	1001	7/7	0.99	0.07	-1.41	104,105,106,106	0
14	ZN	A	1001	1/1	0.98	0.04	-1.81	46,46,46,46	0
14	ZN	I	1002	1/1	0.96	0.06	-1.91	78,78,78,78	0
14	ZN	N	1001	1/1	0.99	0.07	-1.95	56,56,56,56	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å^2)	Q<0.9
14	ZN	P	1001	1/1	0.98	0.05	-1.98	33,33,33,33	0
14	ZN	B	2001	1/1	0.98	0.03	-2.19	66,66,66,66	0
14	ZN	J	2001	1/1	0.99	0.03	-2.35	124,124,124,124	0
14	ZN	W	1001	1/1	0.99	0.11	-4.74	63,63,63,63	0
15	MG	A	1003	1/1	0.98	0.05	-	2,2,2,2	0
15	MG	I	1003	1/1	0.99	0.03	-	2,2,2,2	0

6.5 Other polymers (i)

There are no such residues in this entry.