wwPDB X-ray Structure Validation Summary Report (i)

Feb 6, 2024 – 03:58 AM EST

PDB ID	:	1YNJ
Title	:	Taq RNA polymerase-Sorangicin complex
Authors	:	Campbell, E.A.; Pavlova, O.; Zenkin, N.; Leon, F.; Irschik, H.; Jansen, R.; Severinov, K.; Darst, S.A.
Deposited on	:	2005-01-24
Resolution	:	3.20 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 (i)) were used in the production of this report:

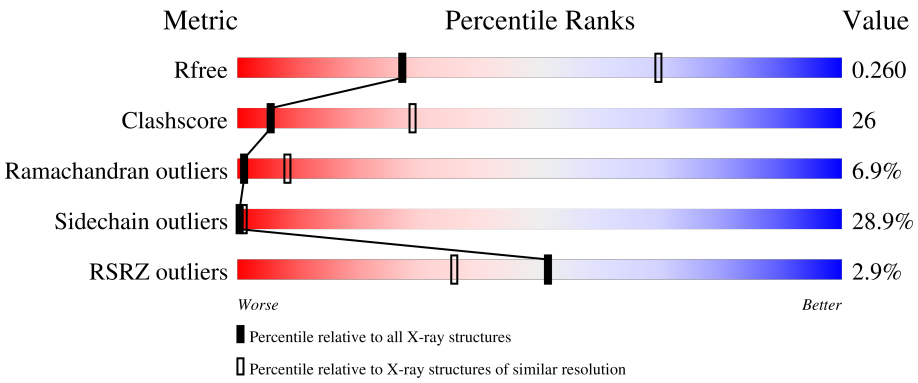
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	314	% 25% 32% 13% • 27%
1	B	314	32% 30% 8% • 28%
2	C	1119	% 39% 42% 15% •
3	D	1524	4% 36% 33% 11% • 19%
3	J	1524	% 7% 7% • 84%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	K	99	2% 42% 35% 17%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 24369 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
			Total	C	N	O	S			
1	A	230	1763	1126	300	334	3	0	0	0
1	B	225	1750	1118	300	329	3	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	C	1114	8578	5431	1516	1607	24	0	0	0

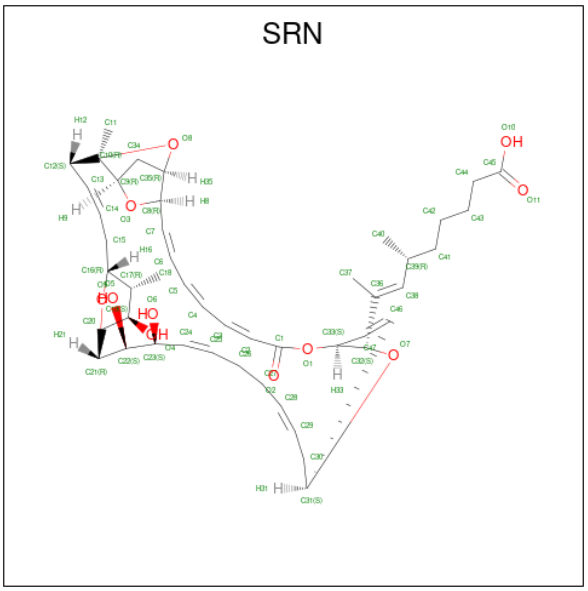
- Molecule 3 is a protein called DNA-directed RNA polymerase beta' chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	D	1238	9602	6065	1703	1798	36	0	0	0
3	J	249	1869	1191	320	356	2	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	K	95	747	476	134	132	5	0	0	0

- Molecule 5 is SORANGICIN A (three-letter code: SRN) (formula: $C_{47}H_{66}O_{11}$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
5	C	1	58	47	11	0	0

- Molecule 6 is ZINC ION (three-letter code: ZN) (formula: Zn).

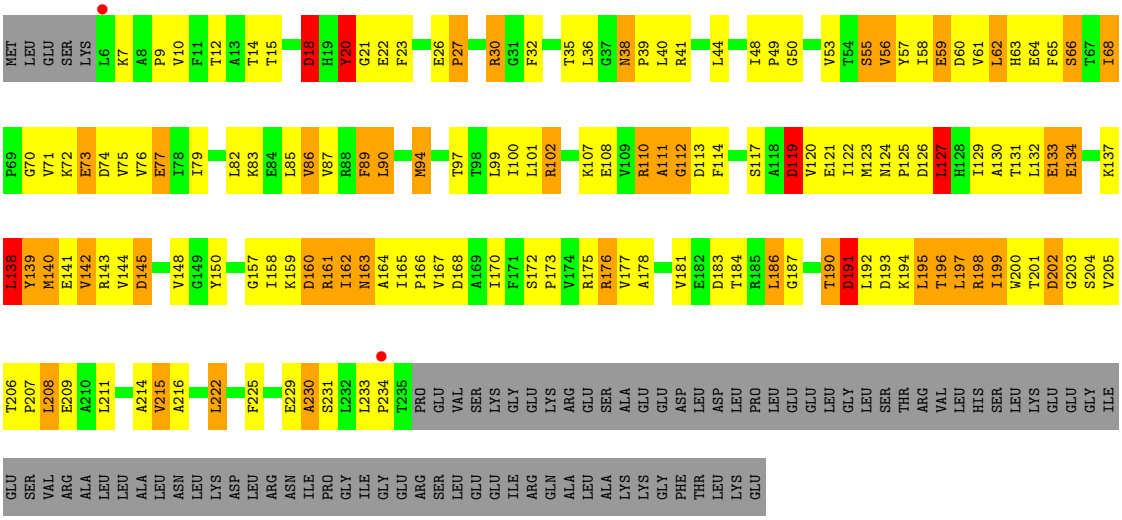
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
6	D	2	2	2	0	0

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.

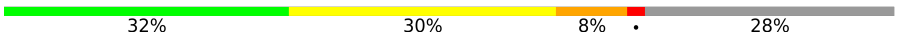
• Molecule 1: DNA-directed RNA polymerase alpha chain

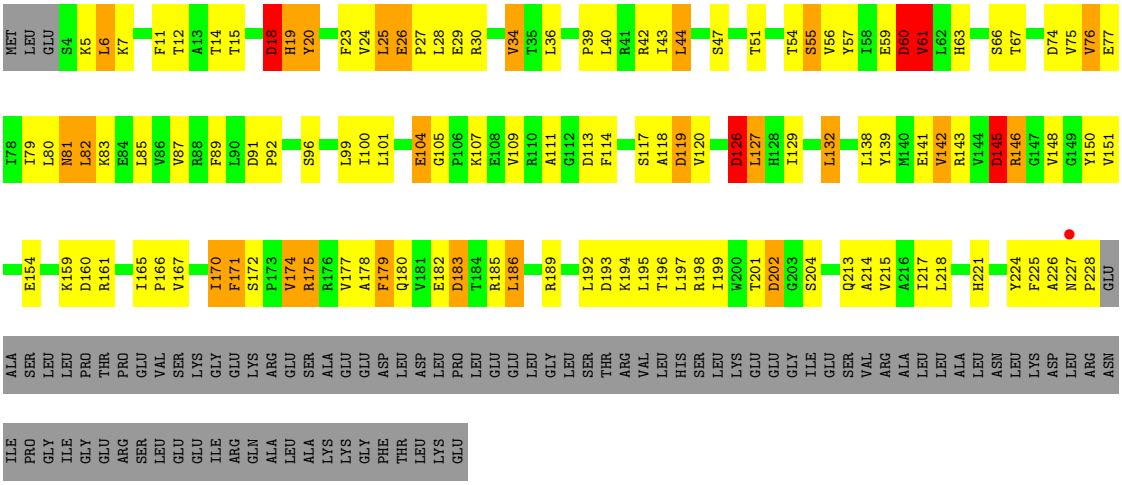
Chain A:



• Molecule 1: DNA-directed RNA polymerase alpha chain

Chain B:





• Molecule 2: DNA-directed RNA polymerase beta chain

Chain C: % 39% 42% 15%

M1 E2 I3 K4 G7 R8 I9 R10 E11 V12 I13 P14 L15 P16 P17 L18 I21 Q22 V23 E24 S25 Y26 K27 K28 A29 L30 Q31 A32 D33 V34 P35 P36 E37 K38 R39 E40 N41 V42 G43 I44 Q45 E50 T51 F52 P53 I54 E55 E56 G57 D58 K61 G62 G63 L64 V65 L66

D67 F68 L69 E70 Y71 R72 I73 G74 D75 P76 P77 F78 S79 Q80 D81 K86 D87 P93 L94 Y95 A96 R97 L98 Q99 L100 I101 H102 K103 D104 T105 G106 K109 E110 D111 E112 V113 F114 L115 L118 P119 T122 E123 D124 G125 S126 F127 I128 I129 A132 D133 R134 V135 I136 V137

S138 Q139 I140 H141 R142 S143 P144 G145 V146 Y147 F148 T149 P150 D151 P152 A153 R154 R157 Y158 I159 A160 S161 P164 L165 I172 D173 L174 E175 V176 E177 A178 S179 G180 V181 V182 T183 M184 K185 V186 N187 K188 F191 P192 L193 V194 L195 L196 L197 R198 V199 L200 G201 Y202 D203 Q204 E205

T206 L207 V208 A213 Y214 G215 D216 L217 V218 Q219 G220 L221 L222 D223 E224 A225 Y226 L227 A228 M229 R230 P231 E232 E233 R237 T240 L241 L242 R243 D246 P247 P248 K249 K250 D251 K252 A253 L254 A255 Y256 L257 P258 G259 L260 L261 A262 D263 P264 K265 R266 Y267 D268 L269 G270 Y275

K276 A277 L281 G282 V283 G284 L285 S286 G287 R288 T289 L290 V291 R292 F293 E294 D295 F298 K299 D300 E301 V302 F303 T306 Y309 A312 A315 G316 V322 D323 D324 H327 L328 G329 N330 R331 R332 I333 R334 T335 V336 G337 E338 L339 M340 Q343 F344 R345 V346 G347 L348

L351 A352 R353 G354 V355 R356 E357 R358 M359 S363 P364 D365 T366 P369 A370 K371 L372 V373 N374 S375 R376 P377 L378 F385 F386 S387 R388 L391 S392 Q393 F394 K395 D396 E397 T398 N399 P400 L401 S402 R405 H406 K407 R408 R409 I410 L413 G416 G417 L418 T419 R422

F425 D429 V430 H431 R432 T433 R437 I438 C439 P440 V441 E442 E445 N448 I449 G450 L451 S454 L455 A456 R460 V461 D462 A463 I467 R468 T469 P470 Y471 R472 R473 V474 K475 V478 V479 V483 V484 Y485 M486 T487 A488 S489 E490 E491 D492 T495 Q498

T501 P502 L503 E504 I508 V513 V514 A515 R516 R517 R518 G519 E520 P521 V522 I523 V524 E528 V529 E530 F531 M532 K537 F540 S541 L542 N543 T544 N545 L546 I547 P548 F549 L550 E551 H552 D553 D554 A555 N556 R557 A558 L559 M560 G561 S562 N563 M564 Q565 T566 V569 P570 L571 I572

R573 A574 Q575 A576 P577 V578 V579 M580 T581 G582 L583 E584 E585 R586 D590 S591 L592 L595 E599 D600 G601 K605 V606 D607 G608 T609 R610 I611 A612 V613 R614 Y615 E616 D617 G618 R619 L620 V621 L625 R626 R627 Y628 A629 R630 S631 N632 Q633 G634 T635 A636 F637 D638 Q639 R640 P641

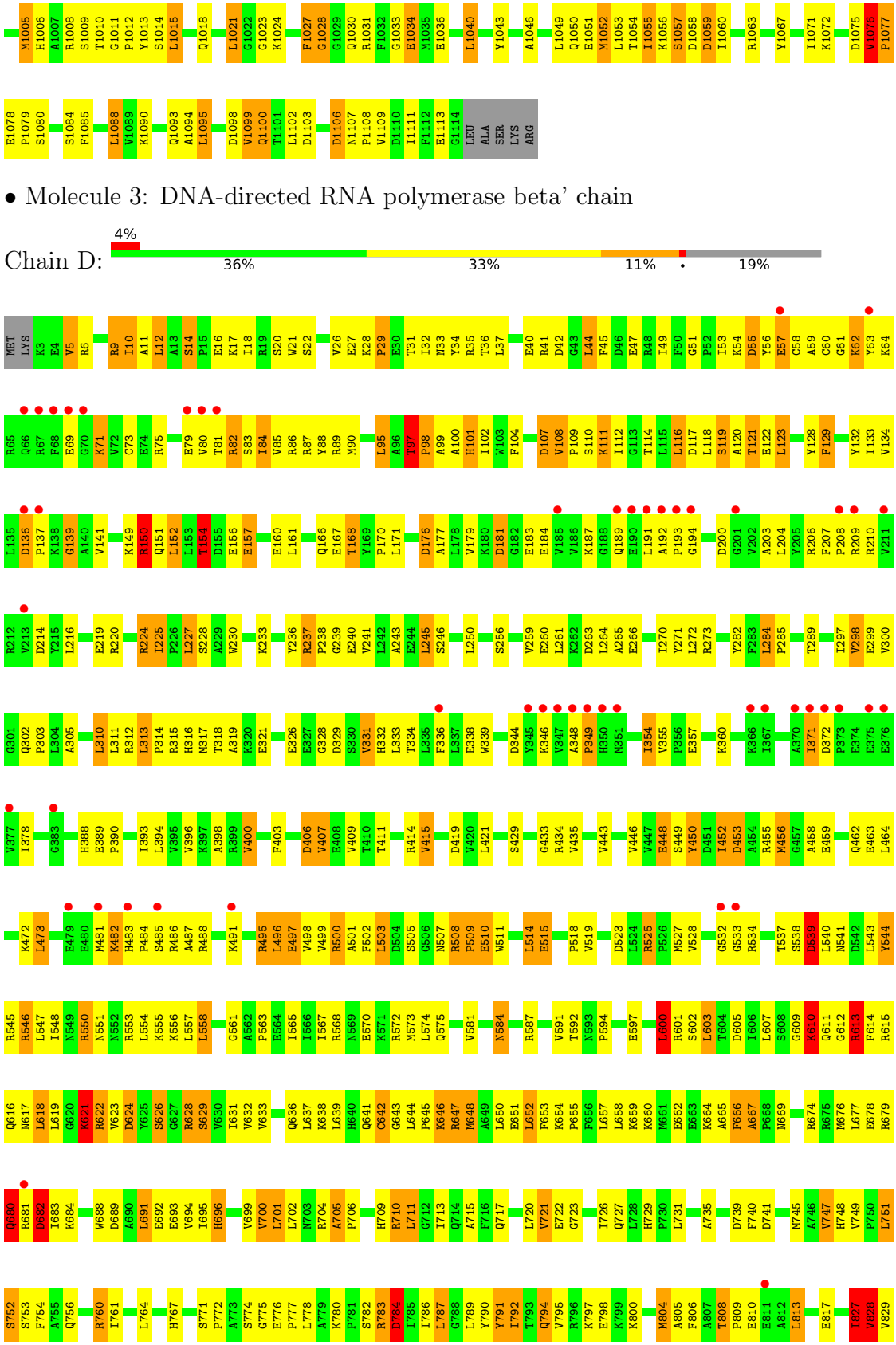
R642 V643 R644 V645 G646 Q647 R648 V649 K650 K651 K652 D653 L654 L655 A656 D657 G658 P659 F665 Q670 N671 V674 A675 I676 M677 P678 F679 D680 G681 V682 N683 F684 E685 I688 V689 I690 S691 E692 E693 K696 R697 D698 F699 Y700 T701 S702 I703 H704 I705 I710 E711 A712 R713 D714

T715 K716 L717 E720 R721 I722 T723 R724 D725 I726 P727 H728 L729 S730 E731 A732 A733 L734 R735 E740 G741 I742 V743 R744 I745 G746 V749 K750 P751 G752 D753 I754 L755 V756 G757 R758 K762 G763 E764 Q765 S768 P769 E770 E771 R772 L773 L774 R775 G779 A782 R783 D784 V785

K786 D787 T788 S789 L790 R791 V792 E796 I799 V800 V801 G802 R803 L804 R805 L806 R807 R808 G809 D810 P811 G812 V813 E814 L815 K816 V819 R820 E821 V822 V823 R824 R831 K832 L833 Q834 K838 L839 A840 N841 R842 H843 G844 N845 K846 K851 I852 L853 P854 V855 E856 D857 M858 P859

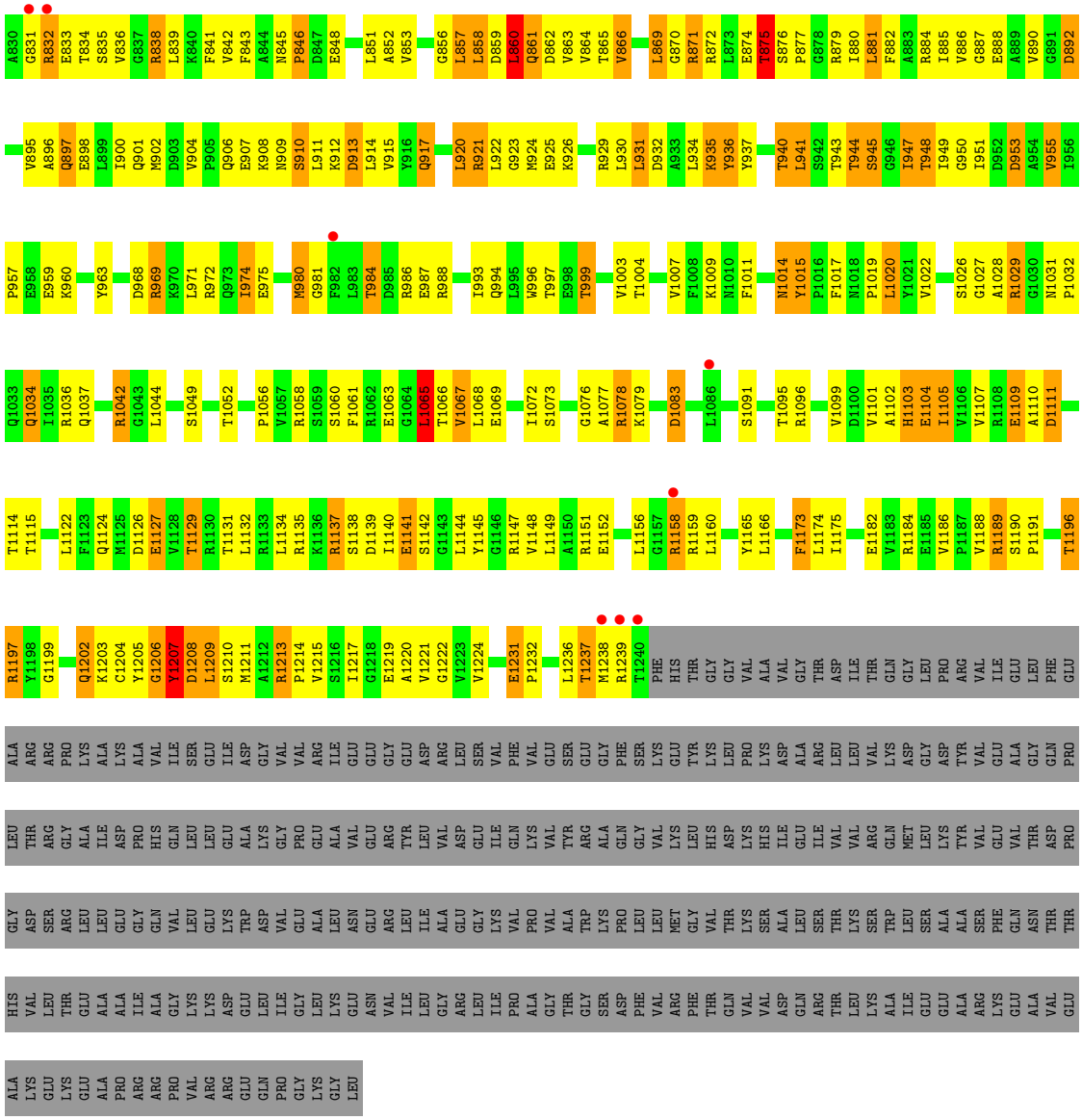
H860 L861 P862 D863 G864 T865 P866 V867 D868 V869 I870 L871 N872 P873 L874 G875 V876 P877 S878 R879 M880 N881 L886 H889 L890 A893 G894 L897 G898 Q899 R900 Y901 I902 S903 P904 V905 F906 D907 G908 A909 T910 E911 P912 E913 I914 K915 E916 L917 L918 F922 N923 L924 Y925 R929

Q930 F934 G935 V936 D937 K938 R939 L944 A945 E948 K949 L950 G951 L952 S958 P959 E960 E961 Q962 L963 K964 E965 I966 F967 D968 L969 G970 K971 V972 V973 L974 Y975 D976 G977 R978 T979 G980 E981 P982 F983 E984 I987 V988 V989 G990 F993 I994 M995 K996 L997 Y998 H999 M1000 V1001



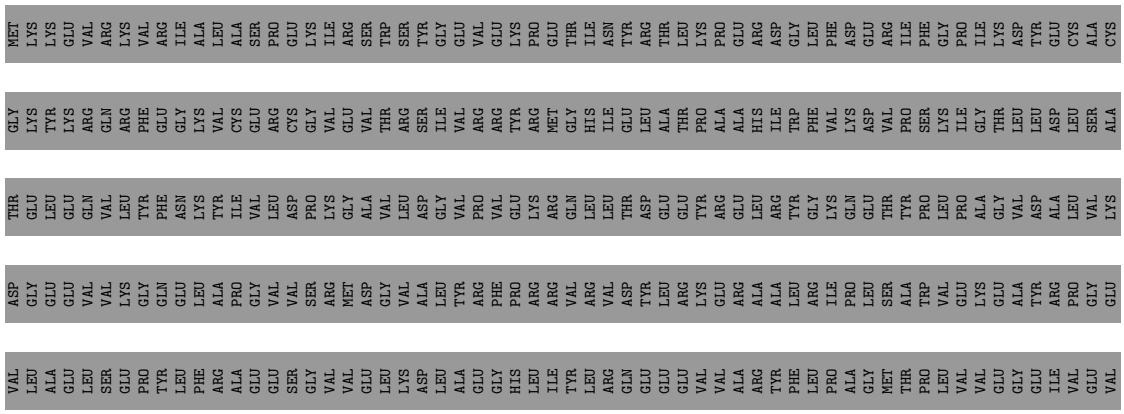
- Molecule 3: DNA-directed RNA polymerase beta' chain

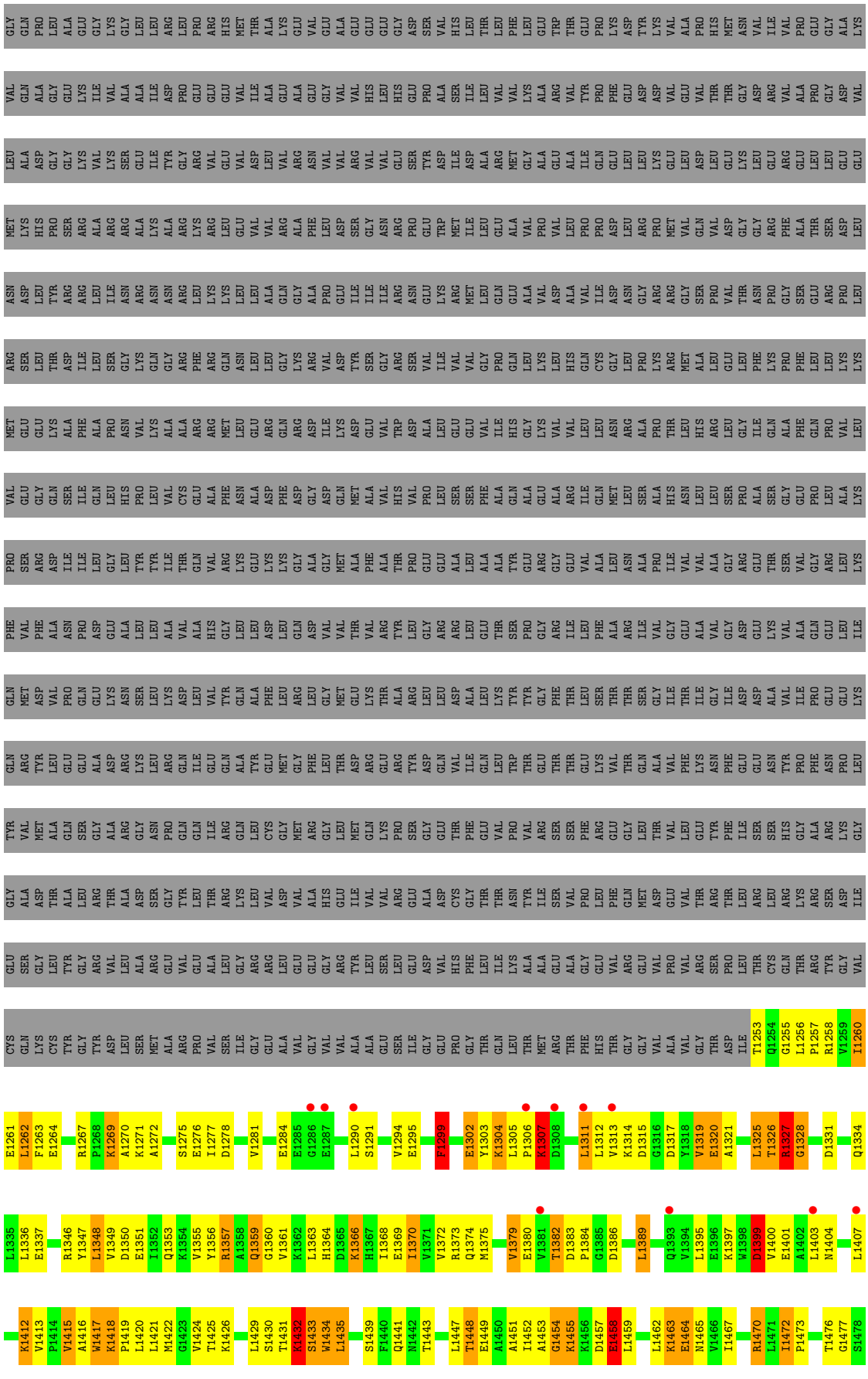
Chain D:

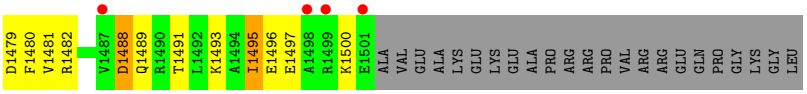


• Molecule 3: DNA-directed RNA polymerase beta' chain

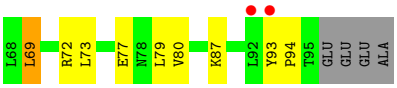
Chain J: % 7% 7% • 84%




GLY GLN PRO LEU ALA GLU GLY LYS GLY LEU LEU ARG LEU PRO ARG HIS MET THR ALA LYS GLU VAL GLU ALA GLU GLU GLU GLY ASP SER VAL HIS LEU THR LEU PHE LEU GLU TRP THR GLU PRO LYS ASP TYR LYS VAL ALA PRO HIS MET ASN VAL ILE VAL PRO GLU GLY ALA LYS
VAL GLN ALA GLY GLU LYS ILE VAL ALA ALA ILE ASP PRO GLU GLU GLU VAL ILE ALA GLU ALA GLU GLY VAL VAL HIS LEU HIS GLU PRO ALA SER ILE LEU VAL ALA ILE ILE TYR GLY ARG VAL GLU VAL ILE ASP ARG VAL VAL LYS ALA ARG VAL TYR PRO PHE GLU ASP ASP VAL GLU VAL ARG VAL
LEU ALA ASP GLY GLY LYS VAL LYS SER GLU ILE TYR ASP ILE ASP ALA ARG MET GLY ALA GLU ALA ILE GLN GLU LEU LYS GLU LEU ASP LEU ARG GLU ARG GLU LEU LEU GLU
MET LYS HIS PRO SER ARG ALA ARG ARG ALA ALA LYS ALA ARG LYS ARG LEU GLU VAL VAL ARG ALA PHE LEU ASP SER GLY ASN ARG PRO GLU TRP MET ILE LEU GLU GLU ALA VAL PRO VAL LEU PRO PRO MET VAL GLN VAL ASP GLY GLY ARG PHE ALA THR SER GLU ASP LEU
ASN ASP LEU TYR ARG ARG LEU ILE ASN ARG ASN ASN ARG LEU LEU ALA GLN GLY LYS LYS LEU LEU GLU GLU ALA ALA PRO GLU ILE ILE ILE ARG ASN GLU LYS ARG LEU GLN GLU ALA VAL ASP VAL ALA ILE ILE ASP ILE GLY SER VAL VAL GLN GLY ASN ARG LEU PRO ARG ASN ARG PRO
ARG SER LEU THR ASP ILE LEU SER GLY LYS GLN GLY ARG PHE ARG GLN LEU LEU GLY LYS ARG VAL ASP TYR SER GLY ARG SER VAL ILE VAL VAL GLY HIS LYS LEU HIS LYS ARG MET ALA LEU GLU GLY LEU PRO LYS ARG MET GLY ARG LEU GLU VAL THR LEU ALA LEU LEU ASN GLY ARG LEU LYS LYS LEU
MET GLU GLU LYS ALA PHE ALA PRO ASN VAL LYS ALA ALA ARG ARG MET LEU GLU ARG GLN ARG ASP ILE LYS ASP GLU VAL TRP ASP ALA LEU GLU GLU VAL LYS VAL ALA ALA PRO THR LEU HIS GLY LYS VAL LEU LEU GLN CYS GLY LEU ASN ARG ALA PRO THR LEU HIS ARG LEU GLY ILE GLN ALA PHE GLN PRO VAL LEU
VAL GLU GLY GLN SER ILE GLN LEU HIS PRO LEU VAL CYS GLU ALA PHE ASN ALA ASP PHE ASP GLY ASP GLN MET ALA VAL HIS VAL PRO LEU SER SER PHE ALA GLN ALA GLU GLU ALA ARG ILE ILE GLN MET LEU ASN ARG ALA SER ALA HIS ASN LEU LEU LEU SER PRO ALA SER GLY GLU PRO LEU LEU LYS
PRO SER ARG ASP ILE ILE LEU GLY LEU TYR TYR ILE THR GLN VAL ARG LYS GLU LYS LYS GLY ALA GLY MET ALA PHE ALA THR PRO GLU GLU ALA LEU ALA ALA GLU GLU VAL VAL ALA LEU ASN ALA PRO ILE VAL VAL ALA GLY ARG GLU THR ARG GLY LEU LYS
PHE VAL PHE ALA ASN PRO ASP GLU ALA LEU LEU ALA VAL ALA HIS GLY LEU LEU ASP LEU GLN VAL VAL THR VAL ARG TYR LEU GLY ARG ARG LEU GLU THR SER PRO GLY ARG ILE LEU GLU GLY ALA ALA ARG ILE VAL GLY GLU ALA ALA VAL GLY GLY ASP GLU LYS VAL GLY ARG GLU LEU ILE
GLN MET ASP VAL PRO GLN GLU LYS ASN SER LEU LYS ASP LEU VAL TYR GLN ALA PHE LEU ARG LEU GLY MET GLU LYS THR ALA ARG LEU ASP LEU LYS TYR TYR GLY PHE THR LEU LEU SER THR THR SER GLY ILE THR ILE ILE ASP ASP ALA VAL GLY ILE GLN LYS LYS GLU
GLN ARG TYR LEU GLU GLU ALA ASP ARG LYS LEU ARG GLN ILE GLU GLN ALA TYR GLU MET GLY PHE LEU THR ASP ARG GLU ARG TYR ASP GLN VAL ILE GLN LEU TRP THR GLU THR GLU LYS VAL THR GLN ALA VAL PHE LYS ASN PHE GLU GLU ASN TYR PRO PHE GLU LEU
TYR VAL MET ALA GLN SER GLY GLY ALA ARG GLY ASN PRO GLN GLN ILE ARG GLN LEU CYS GLY MET ARG GLY LEU MET GLN LYS PRO LEU GLY GLU THR PHE GLU VAL PRO VAL ARG SER SER PHE ARG GLU GLY LEU THR VAL LEU GLU VAL ILE SER SER HIS GLY ALA ARG LYS ARG GLY
GLY ALA ASP THR ALA ALA LEU ARG THR ALA ASP SER GLY TYR LEU VAL ASP VAL ALA HIS GLU ILE VAL VAL ARG GLU ALA ASP CYS GLY THR THR ASN TYR ILE SER VAL PHE GLN MET ASP GLU VAL THR ARG THR PHE LYS ASN PRO VAL LEU ARG LYS ARG SER ASP ILE
GLU SER GLY LEU TYR GLY ARG VAL LEU LEU ALA ARG GLU VAL GLU ALA ARG GLY LEU LEU GLU GLU ARG TYR LEU SER LEU ASP VAL HIS PHE LEU ILE LYS ALA ALA GLU GLU VAL ARG GLU VAL PRO VAL ARG THR ARG THR LEU THR CYS GLN THR ARG TYR GLY VAL
CYS GLN LYS CYS TYR GLY TYR ASP LEU SER MET ALA ARG PRO VAL SER ILE GLY GLU ALA VAL GLY VAL VAL ALA ALA GLU SER ILE GLY GLU PRO GLY THR GLN LEU THR HIS THR GLY GLY VAL ALA VAL GLY THR ASP ILE T1253 Q1254 G1255 L1256 P1257 R1258 V1259 I1260
E1261 L1262 F1263 E1264 R1267 P1268 K1269 A1270 K1271 A1272 S1275 E1276 I1277 D1278 V1281 E1284 E1285 G1286 E1287 L1290 S1291 V1294 E1295 F1299 E1302 Y1303 K1304 L1305 P1306 K1307 D1308 L1311 L1312 V1313 K1314 D1315 G1316 D1317 Y1318 V1319 E1320 A1321 L1325 T1326 R1327 G1328 D1331 Q1334
L1335 L1336 E1337 R1346 Y1347 L1348 V1349 D1350 E1351 I1352 Q1353 K1354 V1355 Y1356 R1357 A1358 Q1359 G1360 V1361 K1362 L1363 H1364 D1365 K1366 H1367 I1368 E1369 I1370 V1371 V1372 R1373 Q1374 M1375 V1379 E1380 V1381 T1382 D1383 P1384 G1385 D1386 L1389 Q1393 V1394 L1395 E1396 K1397 W1398 D1399 V1400 E1401 A1402 L1403 N1404 L1407
K1412 V1413 P1414 V1415 A1416 W1417 K1418 P1419 L1420 L1421 M1422 G1423 V1424 T1425 K1426 L1429 S1430 T1431 K1432 S1433 W1434 L1435 S1439 F1440 Q1441 N1442 T1443 L1447 T1448 E1449 A1450 A1451 I1452 A1453 G1454 K1455 K1456 D1457 E1458 L1459 L1462 K1463 E1464 N1465 V1466 I1467 R1470 L1471 T1472 P1473 T1476 G1477 S1478



• Molecule 4: DNA-directed RNA polymerase omega chain



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	199.19Å 199.19Å 289.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 3.20 39.84 – 3.18	Depositor EDS
% Data completeness (in resolution range)	93.1 (40.00-3.20) 92.3 (39.84-3.18)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.34 (at 3.18Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.281 , 0.346 0.259 , 0.260	Depositor DCC
R_{free} test set	4796 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	69.0	Xtriage
Anisotropy	0.147	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.24 , 44.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.87	EDS
Total number of atoms	24369	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.03% 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 (i)

5.1 Standard geometry (i)

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

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	0.52	0/1798	0.88	10/2453 (0.4%)
1	B	0.52	0/1784	0.87	6/2428 (0.2%)
2	C	0.55	0/8744	0.88	35/11850 (0.3%)
3	D	0.52	0/9772	0.81	27/13234 (0.2%)
3	J	0.49	0/1897	0.81	7/2570 (0.3%)
4	K	0.52	0/762	0.79	1/1029 (0.1%)
All	All	0.53	0/24757	0.84	86/33564 (0.3%)

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	3
3	D	0	2
All	All	0	5

There are no bond length outliers.

The worst 5 of 86 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	680	ASP	CB-CG-OD2	7.16	124.74	118.30
2	C	554	ASP	CB-CG-OD2	6.96	124.56	118.30
3	D	860	LEU	CA-CB-CG	6.92	131.22	115.30
2	C	492	ASP	CB-CG-OD2	6.87	124.48	118.30
1	B	193	ASP	CB-CG-OD2	6.71	124.34	118.30

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	1076	VAL	Peptide
2	C	213	ALA	Peptide
2	C	671	ASN	Peptide
3	D	1207	TYR	Peptide
3	D	21	TRP	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	1763	0	1760	117	0
1	B	1750	0	1775	77	0
2	C	8578	0	8517	538	0
3	D	9602	0	9556	504	0
3	J	1869	0	1876	88	0
4	K	747	0	735	41	0
5	C	58	0	64	5	0
6	D	2	0	0	0	0
All	All	24369	0	24283	1265	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 26.

The worst 5 of 1265 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)
3:D:705:ALA:CB	3:D:706:PRO:HD3	1.72	1.17
3:D:643:GLY:HA3	3:D:727:GLN:H	1.12	1.15
2:C:263:ASP:HB3	2:C:264:PRO:HD3	1.18	1.12
3:D:1110:ALA:HA	3:D:1202:GLN:HB3	1.33	1.09
3:D:551:ASN:O	3:D:555:LYS:HB2	1.51	1.08

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	228/314 (73%)	169 (74%)	45 (20%)	14 (6%)	1 12
1	B	223/314 (71%)	177 (79%)	33 (15%)	13 (6%)	1 13
2	C	1112/1119 (99%)	859 (77%)	168 (15%)	85 (8%)	1 7
3	D	1236/1524 (81%)	971 (79%)	186 (15%)	79 (6%)	1 10
3	J	247/1524 (16%)	190 (77%)	39 (16%)	18 (7%)	1 7
4	K	93/99 (94%)	66 (71%)	18 (19%)	9 (10%)	0 3
All	All	3139/4894 (64%)	2432 (78%)	489 (16%)	218 (7%)	1 8

5 of 218 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	20	TYR
1	A	111	ALA
1	A	139	TYR
1	A	187	GLY
1	A	203	GLY

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	189/270 (70%)	135 (71%)	54 (29%)	0 1
1	B	191/270 (71%)	138 (72%)	53 (28%)	0 1
2	C	889/936 (95%)	633 (71%)	256 (29%)	0 1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	D	992/1281 (77%)	715 (72%)	277 (28%)	0	1
3	J	191/1281 (15%)	123 (64%)	68 (36%)	0	0
4	K	75/88 (85%)	53 (71%)	22 (29%)	0	1
All	All	2527/4126 (61%)	1797 (71%)	730 (29%)	0	1

5 of 730 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	488	ARG
3	D	930	LEU
3	D	534	ARG
3	D	486	ARG
3	D	700	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 51 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	1030	GLN
3	D	680	GLN
4	K	29	GLN
3	D	33	ASN
3	D	616	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 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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	$\#\|Z\| > 2$
5	SRN	C	1120	-	60,62,62	1.34	4 (6%)	62,84,84	1.63	10 (16%)

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
5	SRN	C	1120	-	-	10/52/105/105	0/4/5/5

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	1120	SRN	O1-C1	6.91	1.48	1.34
5	C	1120	SRN	C2-C1	-4.71	1.37	1.48
5	C	1120	SRN	C5-C6	-3.50	1.34	1.44
5	C	1120	SRN	C3-C4	-3.12	1.35	1.44

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	C	1120	SRN	O1-C1-C2	5.25	123.29	111.38
5	C	1120	SRN	O1-C33-C47	-3.80	100.97	108.85
5	C	1120	SRN	C34-C9-C10	-3.53	106.16	114.05
5	C	1120	SRN	C35-O8-C12	3.22	118.53	112.81
5	C	1120	SRN	C8-C7-C6	-3.18	119.38	125.61

There are no chirality outliers.

5 of 10 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	C	1120	SRN	O7-C32-C36-C37
5	C	1120	SRN	O2-C1-O1-C33
5	C	1120	SRN	C42-C43-C44-C45
5	C	1120	SRN	C2-C1-O1-C33
5	C	1120	SRN	C39-C41-C42-C43

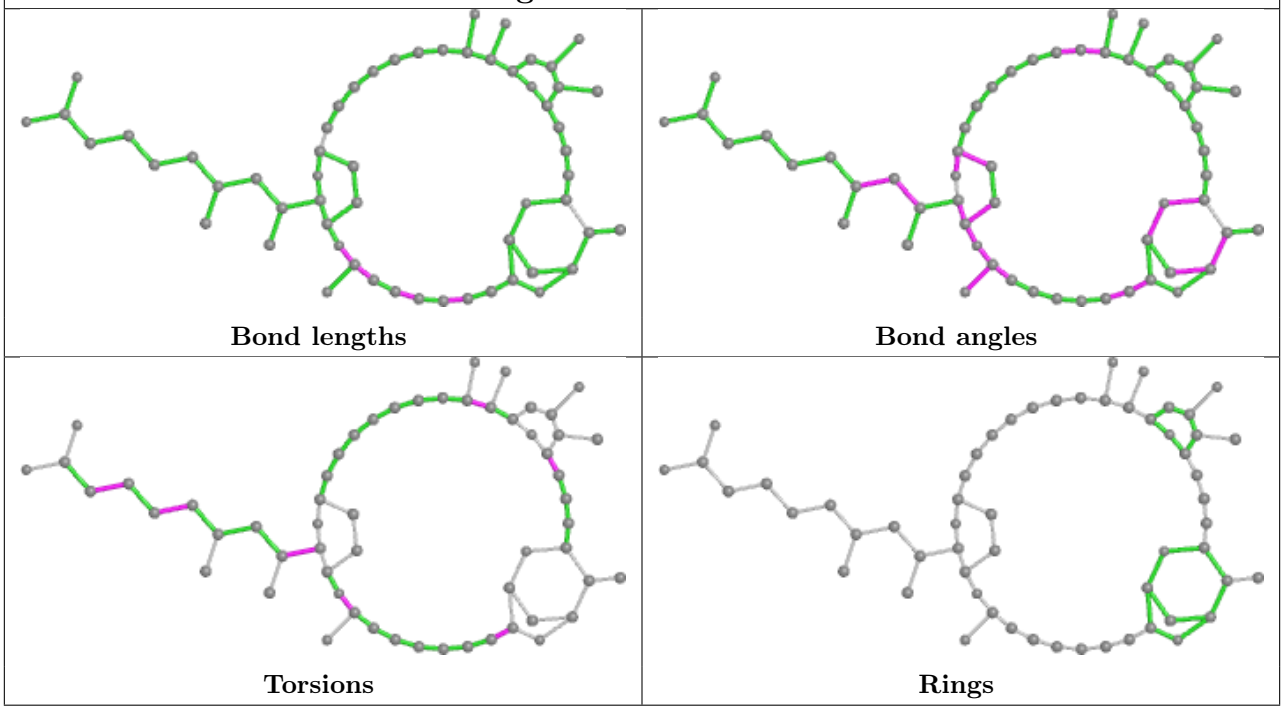
There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	C	1120	SRN	5	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 SRN C 1120



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 (i)

6.1 Protein, DNA and RNA chains (i)

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, 95th 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	230/314 (73%)	-0.27	2 (0%)	84	75	22, 43, 76, 112	0
1	B	225/314 (71%)	-0.33	1 (0%)	92	89	24, 51, 81, 86	0
2	C	1114/1119 (99%)	-0.28	13 (1%)	79	67	19, 50, 90, 115	0
3	D	1238/1524 (81%)	-0.10	59 (4%)	30	18	18, 49, 94, 139	0
3	J	249/1524 (16%)	0.02	15 (6%)	21	12	24, 51, 125, 139	0
4	K	95/99 (95%)	-0.24	2 (2%)	63	49	27, 64, 149, 162	0
All	All	3151/4894 (64%)	-0.19	92 (2%)	51	36	18, 50, 95, 162	0

The worst 5 of 92 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	D	191	LEU	8.9
3	D	193	PRO	7.3
3	D	346	LYS	5.3
3	D	1240	THR	5.3
3	D	373	PRO	5.0

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 monosaccharides in this entry.

6.4 Ligands (i)

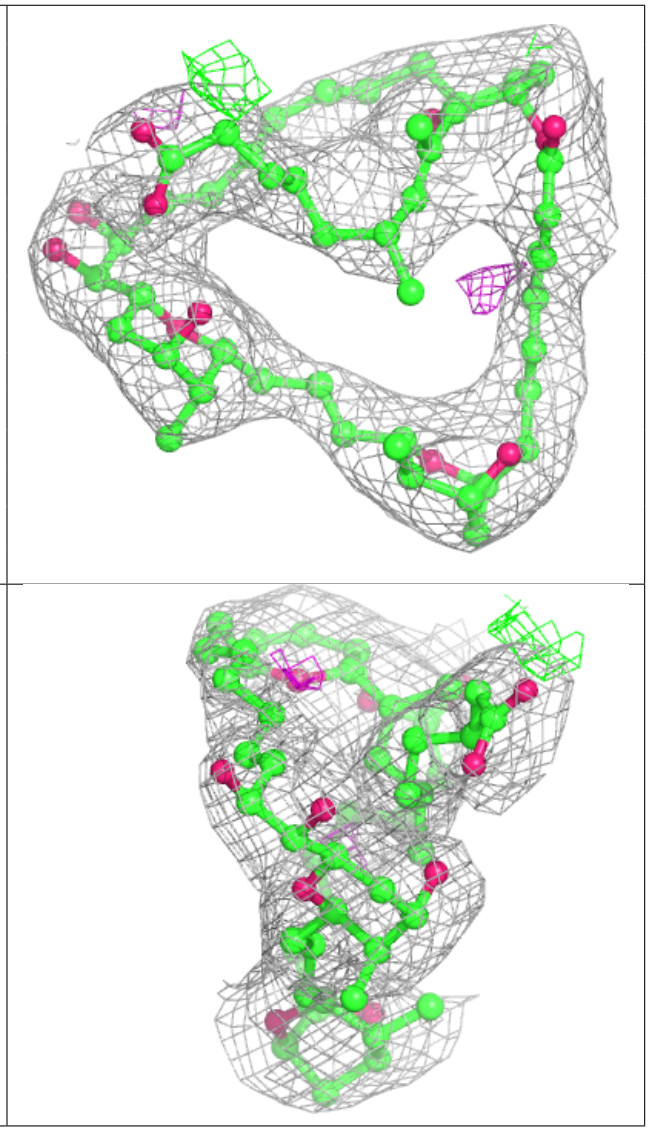
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	SRN	C	1120	58/58	0.92	0.27	33,45,56,59	0
6	ZN	D	1526	1/1	0.93	0.13	64,64,64,64	0
6	ZN	D	1525	1/1	0.96	0.03	77,77,77,77	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 SRN C 1120:

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.