wwPDB X-ray Structure Validation Summary Report (i)

Oct 18, 2023 – 08:26 PM EDT

PDB ID	:	2E2I
Title	:	RNA polymerase II elongation complex in 5 mM Mg+2 with 2'-dGTP
Authors	:	Wang, D.; Bushnell, D.A.; Westover, K.D.; Kaplan, C.D.; Kornberg, R.D.
Deposited on	:	2006-11-14
Resolution	:	3.41 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:

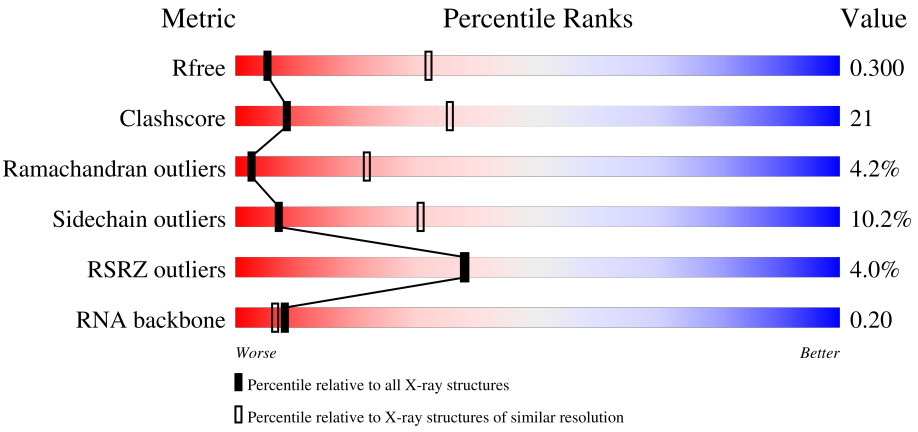
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 3.41 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.



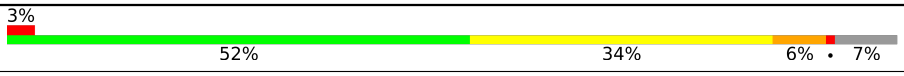
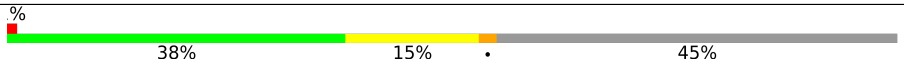
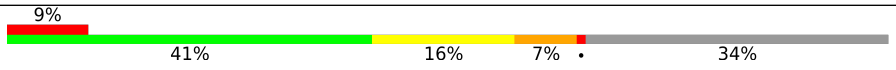
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1486 (3.50-3.34)
Clashscore	141614	1572 (3.50-3.34)
Ramachandran outliers	138981	1534 (3.50-3.34)
Sidechain outliers	138945	1535 (3.50-3.34)
RSRZ outliers	127900	1395 (3.50-3.34)
RNA backbone	3102	1012 (3.88-2.96)

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	R	10	10% 10% 20% 50% 20%
2	T	28	4% 7% 75% 18%
3	N	14	7% 7% 64% 29%
4	A	1733	4% 51% 26% 19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	B	1224	
6	C	318	
7	E	215	
8	F	155	
9	H	146	
10	I	122	
11	J	70	
12	K	120	
13	L	70	

2 Entry composition (i)

There are 16 unique types of molecules in this entry. The entry contains 29722 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 RNA chain called 5'-R(*AP*UP*CP*GP*AP*GP*AP*GP*GP*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	R	10	Total	C	N	O	P	0	0	0
			216	98	45	64	9			

- Molecule 2 is a DNA chain called 28-MER DNA template strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	T	28	Total	C	N	O	P	0	0	0
			564	270	102	165	27			

- Molecule 3 is a DNA chain called 5'-D(*CP*TP*GP*CP*TP*TP*AP*TP*CP*GP*GP*TP*AP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	14	Total	C	N	O	P	0	0	0
			284	137	49	85	13			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	A	1411	Total	C	N	O	S	0	0	0
			11098	6989	1944	2104	61			

- Molecule 5 is a protein called DNA-directed RNA polymerase II 140 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	B	1143	Total	C	N	O	S	0	0	0
			9092	5753	1595	1688	56			

- Molecule 6 is a protein called DNA-directed RNA polymerase II 45 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	C	266	Total	C	N	O	S	0	0	0
			2095	1317	348	417	13			

- Molecule 7 is a protein called DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide.

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

- Molecule 8 is a protein called DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	F	86	Total	C	N	O	S	0	0	0
			697	445	118	131	3			

- Molecule 9 is a protein called DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
9	H	134	Total	C	N	O	S	0	0	0
			1076	678	181	212	5			

- Molecule 10 is a protein called DNA-directed RNA polymerase II subunit 9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
10	I	119	Total	C	N	O	S	0	0	0
			971	596	179	186	10			

- Molecule 11 is a protein called DNA-directed RNA polymerases I/II/III subunit 10.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
11	J	65	Total	C	N	O	S	0	0	0
			532	339	93	94	6			

- Molecule 12 is a protein called DNA-directed RNA polymerase II 13.6 kDa polypeptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
12	K	114	Total	C	N	O	S	0	0	0
			919	590	156	171	2			

- Molecule 13 is a protein called DNA-directed RNA polymerases I, II, and III 7.7 kDa polypep-

tide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
13	L	46	Total	C	N	O	S	0	0	0
			363	224	72	63	4			

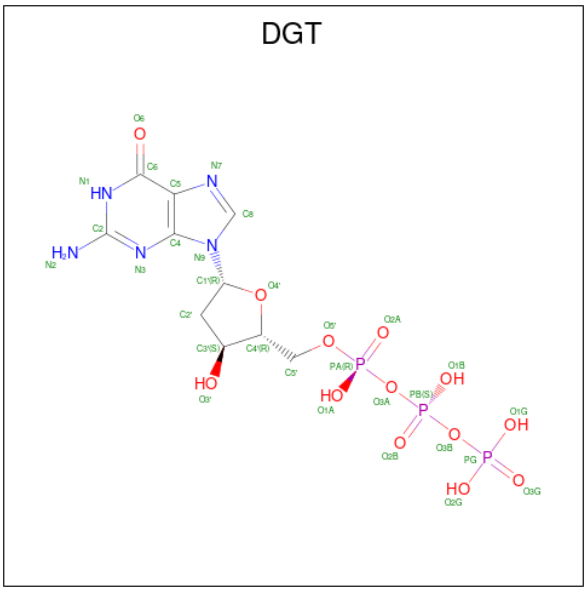
- Molecule 14 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
14	A	2	Total	Zn	0	0
			2	2		
14	B	1	Total	Zn	0	0
			1	1		
14	C	1	Total	Zn	0	0
			1	1		
14	I	2	Total	Zn	0	0
			2	2		
14	J	1	Total	Zn	0	0
			1	1		
14	L	1	Total	Zn	0	0
			1	1		

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

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

- Molecule 16 is 2'-DEOXYGUANOSINE-5'-TRIPHOSPHATE (three-letter code: DGT) (formula: $C_{10}H_{16}N_5O_{13}P_3$).



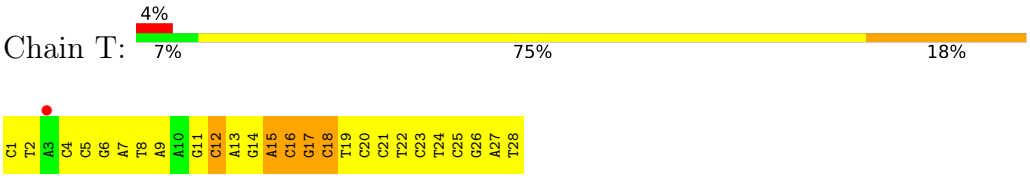
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
16	B	1	Total	C	N	O	P	0	1
			62	20	10	26	6		

3 Residue-property plots (i)

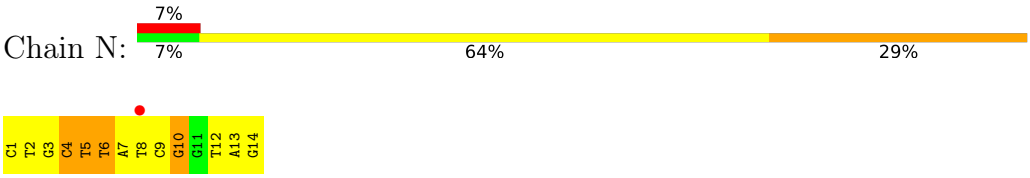
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: 5'-R(*AP*UP*CP*GP*AP*GP*AP*GP*GP*A)-3'

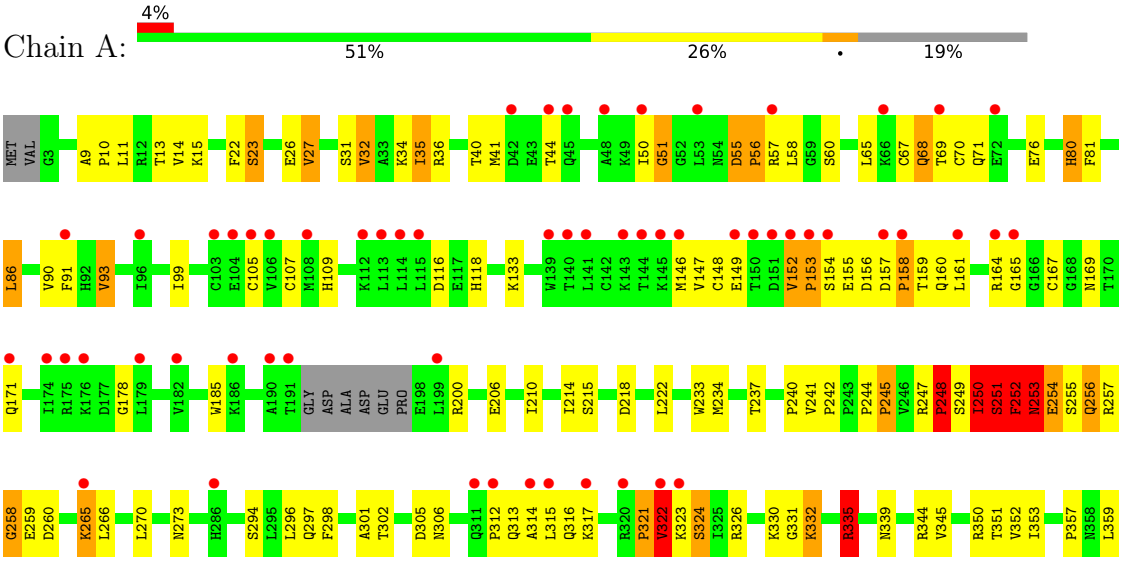
• Molecule 2: 28-MER DNA template strand

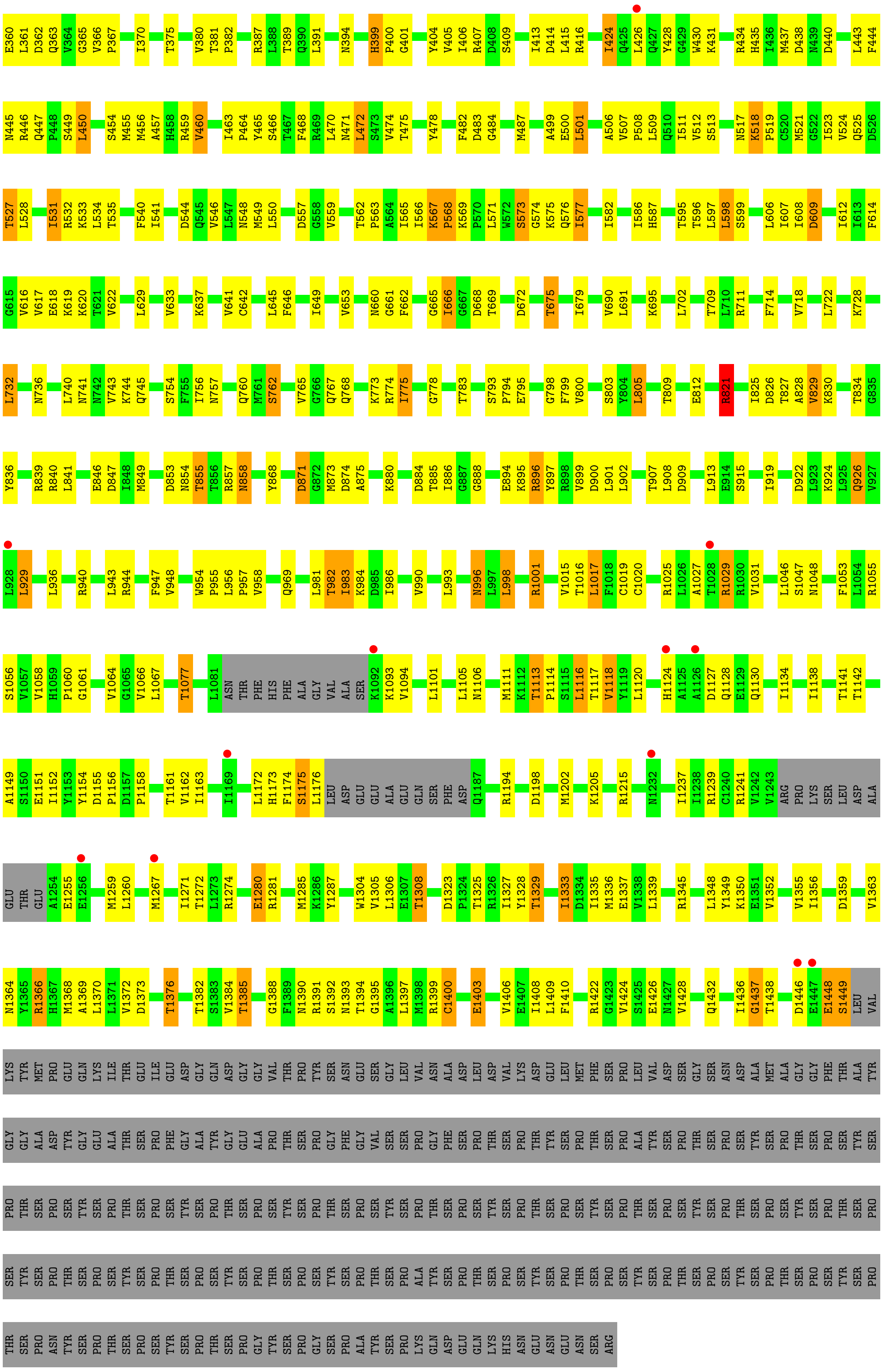


• Molecule 3: 5'-D(*CP*TP*GP*CP*TP*TP*AP*TP*CP*GP*GP*TP*AP*G)-3'



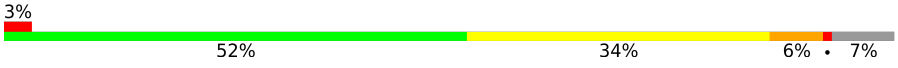
• Molecule 4: DNA-directed RNA polymerase II largest subunit

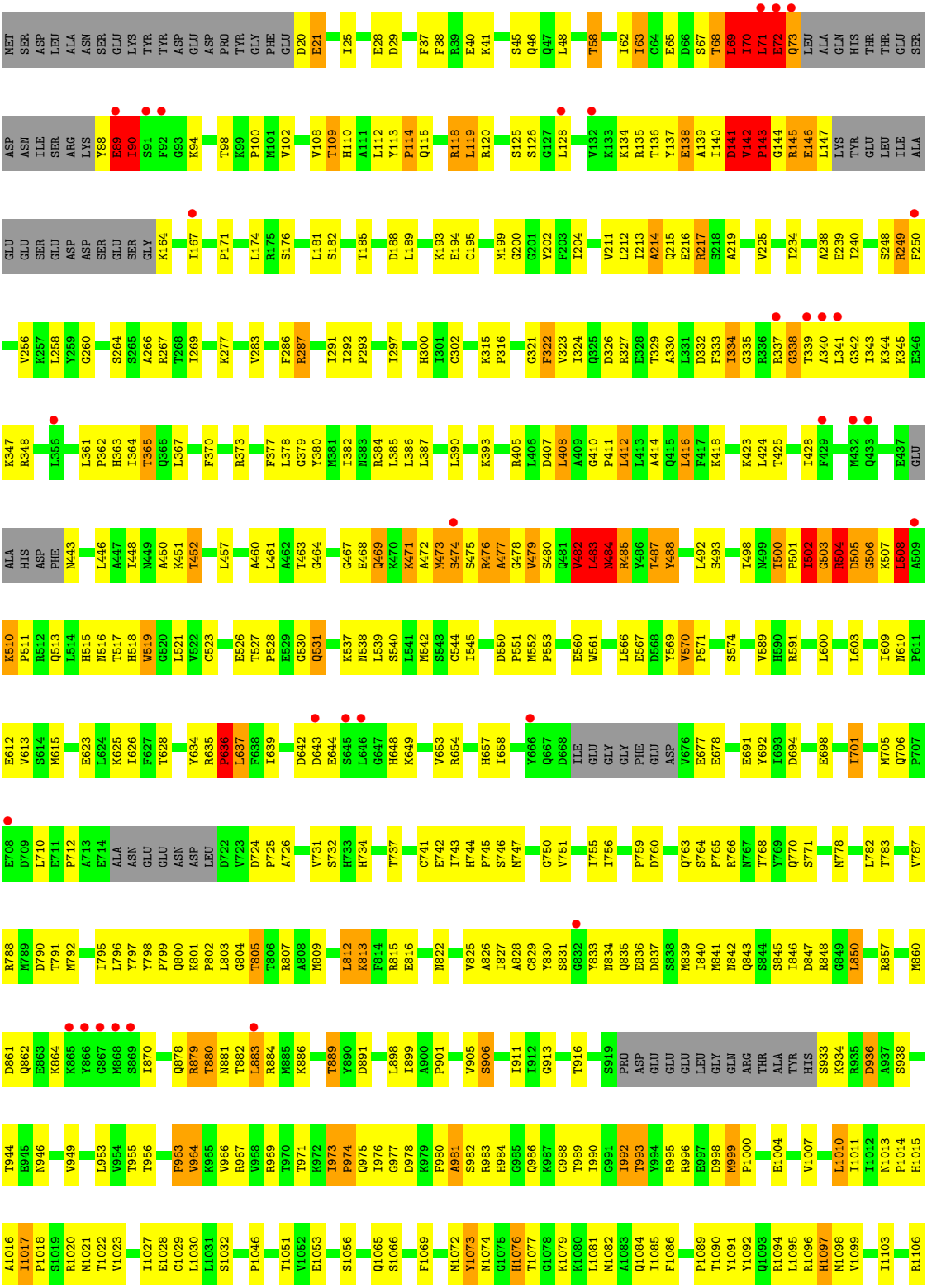


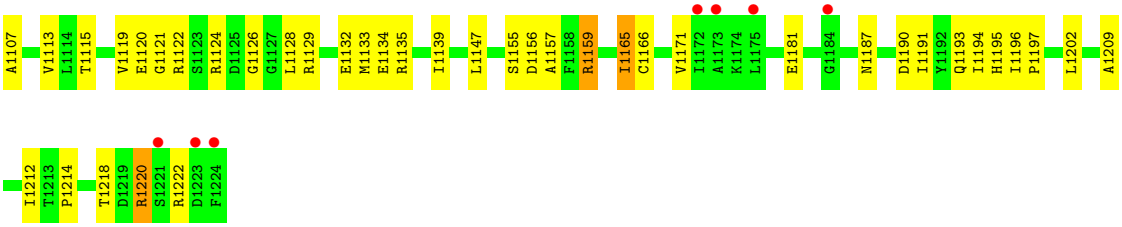

E360 L361 D362 Q363 V364 G365 V366 P367 I370 T375 V380 T381 P382 R387 L388 T389 Q390 L391 N394 H399 P400 G401 Y404 V405 I406 R407 D408 S409 I413 D414 L415 R416 I424 Q425 L426 Q427 Y428 G429 W430 K431 R434 H435 I436 M437 D438 N439 D440 L443 F444
N445 R446 Q447 P448 S449 L450 S454 M455 M456 A457 H458 R459 V460 I463 P464 Y465 S466 T467 F468 R469 L470 N471 L472 S473 V474 T475 Y478 F482 D483 G484 M487 A499 E500 L501 A506 V507 P508 L509 Q510 I511 V512 S513 N517 K518 P519 C520 M521 G522 I523 V524 Q525 D526
T527 L528 I531 R532 K533 L534 T535 F540 I541 D544 Q545 V546 L547 N548 M549 L550 D557 G558 V559 T562 P563 A564 I565 I566 K567 P568 K569 P570 L571 W572 S573 G574 K575 Q576 I577 I582 I586 H587 T595 T596 L597 L598 S599 L606 I607 I608 D609 I612 I613 F614
G615 V616 V617 E618 K619 K620 T621 V622 L629 V633 K637 V641 C642 L645 F646 I649 V653 N660 G661 F662 G665 I666 G667 D668 T669 D672 T675 I679 V690 L691 K695 L702 T709 L710 R711 F714 V718 L722 K728
L732 N736 L740 N741 N742 V743 K744 Q745 S754 F755 I756 N757 Q760 M761 S762 V765 G766 Q767 Q768 K773 R774 I775 G778 T783 S793 P794 E795 G798 F799 V800 S803 Y804 L805 T809 E812 R821 I825 D826 T827 A828 V829 K830 T834 G835
Y836 R839 R840 L841 E846 D847 I848 M849 D853 N854 T855 T856 R857 N858 Y868 D871 G872 M873 D874 A875 K880 D884 T885 I886 G887 G888 E894 K895 R896 Y897 R898 V899 D900 L901 L902 T907 L908 D909 L913 E914 S915 I919 D922 L923 K924 L925 Q926 V927
L928 L929 L936 R940 L943 R944 F947 V948 W954 P955 L956 P957 V958 Q969 L981 T982 I983 K984 D985 I986 V990 L993 N996 L997 L998 R1001 V1015 T1016 L1017 F1018 C1019 C1020 R1025 L1026 A1027 T1028 R1029 R1030 V1031 L1046 S1047 N1048 F1053 L1054 R1055
S1056 V1057 V1058 H1059 P1060 G1061 V1064 G1065 V1066 L1067 T1077 L1081 ASN THR PHE HIS PHE ALA GLY VAL ALA SER K1092 K1093 V1094 L1101 L1105 N1106 M1111 K1112 T1113 P1114 S1115 L1116 T1117 V1118 Y1119 L1120 H1124 A1125 A1126 D1127 Q1128 E1129 Q1130 I1134 I1138 T1141 T1142
A1149 S1150 E1151 I1152 Y1153 Y1154 D1155 P1156 D1157 P1158 T1161 V1162 I1163 I1169 L1172 H1173 F1174 S1175 L1176 LEU ASP GLU GLU ALA GLU GLN SER PHE ASP Q1187 R1194 D1198 M1202 K1205 R1215 N1232 I1237 I1238 R1239 C1240 R1241 V1242 V1243 ARG PRO LYS SER LEU ASP ALA
GLU THR GLU A1254 E1255 E1256 M1259 L1260 M1267 I1271 T1272 L1273 R1274 E1280 R1281 M1285 K1286 Y1287 W1304 V1305 L1306 E1307 T1308 D1323 P1324 T1325 R1326 I1327 Y1328 T1329 I1333 D1334 I1335 M1336 E1337 V1338 L1339 R1345 L1348 Y1349 K1350 E1351 V1352 V1355 I1356 D1359 V1363
N1364 Y1365 R1366 H1367 M1368 A1369 L1370 L1371 V1372 D1373 T1376 T1382 S1383 V1384 T1385 G1388 F1389 N1390 R1391 S1392 N1393 T1394 G1395 A1396 L1397 M1398 R1399 C1400 E1403 V1406 E1407 I1408 L1409 F1410 R1422 G1423 V1424 S1425 E1426 N1427 V1428 Q1432 I1436 G1437 T1438 D1446 E1447 E1448 S1449 LEU VAL
LYS TYR MET PRO GLU GLN LYS ILE THR GLU ILE GLU ASP GLY GLN ASP GLY GLY VAL THR PRO TYR SER ASN GLU SER GLY LEU VAL ASN ALA ASP LEU ASP VAL LYS ASP GLU LEU MET PHE SER PRO LEU VAL ASP SER GLY SER ASN ASP ALA MET ALA GLY GLY PHE THR ALA TYR
GLY GLY ALA ASP TYR GLY GLU ALA THR SER PRO PHE GLY ALA TYR GLY GLU ALA PRO THR SER PRO GLY PHE GLY VAL SER SER PRO GLY PHE SER PRO THR SER PRO THR TYR SER PRO THR SER PRO ALA TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER
PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO
SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO ALA TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO SER TYR SER PRO
THR SER PRO ASN TYR SER PRO THR SER PRO SER TYR SER PRO THR SER PRO GLY TYR SER PRO GLY SER PRO ALA TYR SER PRO LYS GLN ASP GLU GLN LYS HIS ASN GLU ASN GLU ASN SER ARG

• Molecule 5: DNA-directed RNA polymerase II 140 kDa polypeptide

Chain B:

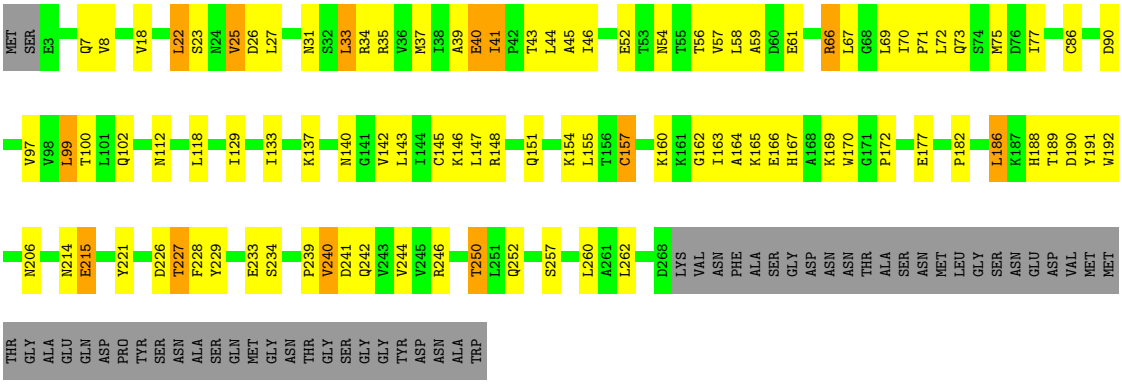






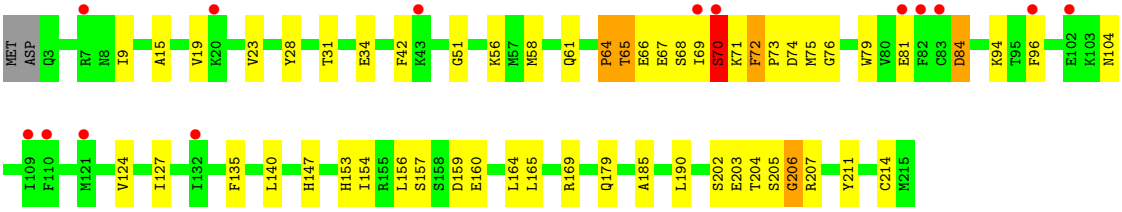
• Molecule 6: DNA-directed RNA polymerase II 45 kDa polypeptide

Chain C:



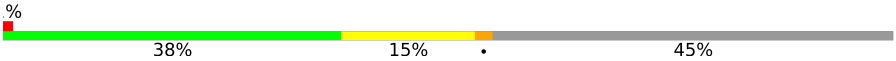
• Molecule 7: DNA-directed RNA polymerases I, II, and III 27 kDa polypeptide

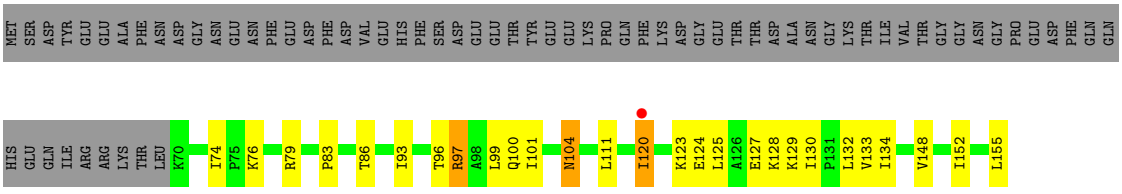
Chain E:



• Molecule 8: DNA-directed RNA polymerases I, II, and III 23 kDa polypeptide

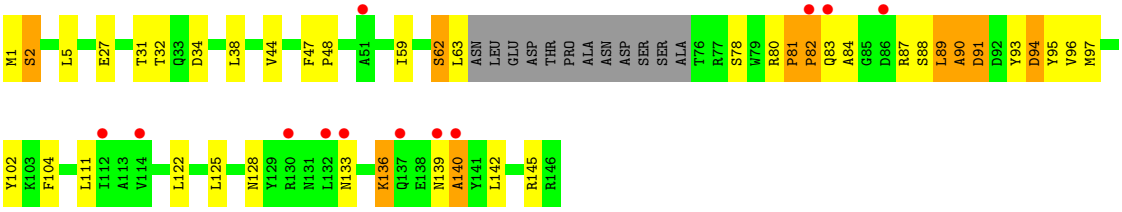
Chain F:



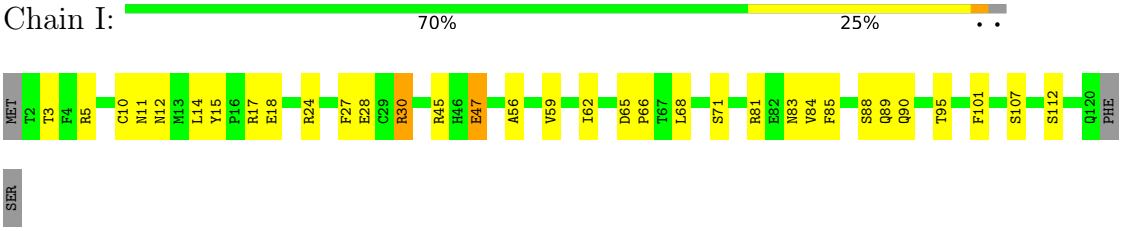


• Molecule 9: DNA-directed RNA polymerases I, II, and III 14.5 kDa polypeptide

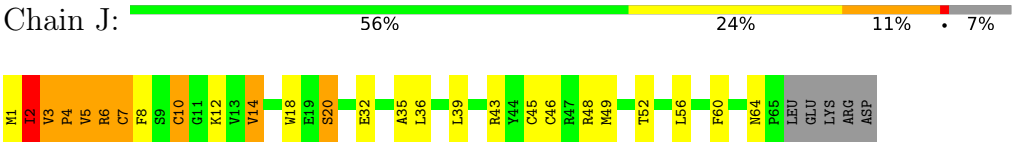
Chain H:



• Molecule 10: DNA-directed RNA polymerase II subunit 9

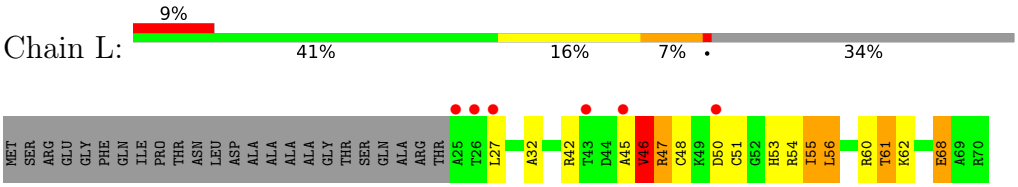


• Molecule 11: DNA-directed RNA polymerases I/II/III subunit 10



• Molecule 12: DNA-directed RNA polymerase II 13.6 kDa polypeptide

• Molecule 13: DNA-directed RNA polymerases I, II, and III 7.7 kDa polypeptide



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	168.68Å 223.52Å 193.94Å 90.00° 100.54° 90.00°	Depositor
Resolution (Å)	50.00 – 3.41 49.77 – 3.41	Depositor EDS
% Data completeness (in resolution range)	98.0 (50.00-3.41) 98.0 (49.77-3.41)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.14	Depositor
$< I/\sigma(I) >$ [1]	3.01 (at 3.40Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.266 , 0.316 0.257 , 0.300	Depositor DCC
R_{free} test set	4732 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	92.7	Xtriage
Anisotropy	0.155	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.26 , 72.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	29722	wwPDB-VP
Average B, all atoms (Å^2)	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.72% 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: ZN, DGT, 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	R	1.15	0/243	1.89	10/378 (2.6%)
2	T	0.87	0/631	1.67	12/970 (1.2%)
3	N	0.81	0/317	1.53	9/488 (1.8%)
4	A	0.65	3/11294 (0.0%)	0.74	3/15270 (0.0%)
5	B	0.80	4/9268 (0.0%)	0.87	9/12496 (0.1%)
6	C	0.74	1/2133 (0.0%)	0.79	1/2891 (0.0%)
7	E	0.51	0/1780	0.64	0/2395
8	F	0.55	0/709	0.69	0/956
9	H	0.51	0/1094	0.66	0/1480
10	I	0.60	0/989	0.66	0/1331
11	J	0.85	1/541 (0.2%)	0.91	0/727
12	K	0.72	1/937 (0.1%)	0.72	0/1265
13	L	0.81	2/365 (0.5%)	0.91	0/485
All	All	0.71	12/30301 (0.0%)	0.84	44/41132 (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
4	A	0	7
5	B	0	11
7	E	0	1
11	J	0	2
All	All	0	21

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	J	10	CYS	CB-SG	7.21	1.94	1.82

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	C	86	CYS	CB-SG	-6.58	1.71	1.82
5	B	302	CYS	CB-SG	-6.13	1.71	1.82
12	K	91	CYS	CB-SG	-6.02	1.72	1.82
5	B	147	LEU	C-O	5.98	1.34	1.23

The worst 5 of 44 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	T	18	DC	O4'-C4'-C3'	-16.51	96.09	106.00
2	T	28	DT	C1'-O4'-C4'	-12.85	97.25	110.10
2	T	18	DC	C4'-C3'-C2'	-12.19	92.13	103.10
1	R	9	G	C5'-C4'-C3'	-9.98	100.04	116.00
2	T	18	DC	C1'-O4'-C4'	-8.74	101.36	110.10

There are no chirality outliers.

5 of 21 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	A	248	PRO	Peptide
4	A	250	ILE	Peptide
4	A	251	SER	Peptide
4	A	252	PHE	Peptide
4	A	254	GLU	Peptide

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	R	216	0	109	13	0
2	T	564	0	316	88	0
3	N	284	0	161	19	0
4	A	11098	0	11174	428	0
5	B	9092	0	9135	504	0
6	C	2095	0	2051	95	0
7	E	1744	0	1772	46	0
8	F	697	0	720	19	0
9	H	1076	0	1052	32	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
10	I	971	0	928	16	0
11	J	532	0	542	48	0
12	K	919	0	929	37	0
13	L	363	0	388	6	0
14	A	2	0	0	0	0
14	B	1	0	0	0	0
14	C	1	0	0	0	0
14	I	2	0	0	0	0
14	J	1	0	0	0	0
14	L	1	0	0	0	0
15	A	1	0	0	0	0
16	B	62	0	24	2	0
All	All	29722	0	29301	1225	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 21.

The worst 5 of 1225 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)
5:B:69:LEU:HG	5:B:70:ILE:CG2	1.32	1.55
2:T:19:DT:H2'	2:T:20:DC:C6	1.41	1.51
5:B:142:VAL:CG1	5:B:144:GLY:HA3	1.43	1.45
5:B:142:VAL:HG13	5:B:144:GLY:CA	1.57	1.34
5:B:69:LEU:CA	5:B:70:ILE:HB	1.58	1.28

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
4	A	1401/1733 (81%)	1186 (85%)	160 (11%)	55 (4%)	3 23
5	B	1129/1224 (92%)	938 (83%)	133 (12%)	58 (5%)	2 17
6	C	264/318 (83%)	226 (86%)	33 (12%)	5 (2%)	8 37
7	E	211/215 (98%)	176 (83%)	29 (14%)	6 (3%)	5 30
8	F	84/155 (54%)	74 (88%)	7 (8%)	3 (4%)	3 25
9	H	130/146 (89%)	104 (80%)	21 (16%)	5 (4%)	3 23
10	I	117/122 (96%)	95 (81%)	19 (16%)	3 (3%)	5 31
11	J	63/70 (90%)	55 (87%)	4 (6%)	4 (6%)	1 12
12	K	112/120 (93%)	102 (91%)	8 (7%)	2 (2%)	8 38
13	L	44/70 (63%)	30 (68%)	7 (16%)	7 (16%)	0 2
All	All	3555/4173 (85%)	2986 (84%)	421 (12%)	148 (4%)	3 22

5 of 148 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	A	50	ILE
4	A	55	ASP
4	A	56	PRO
4	A	250	ILE
4	A	312	PRO

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
4	A	1233/1520 (81%)	1121 (91%)	112 (9%)	9 35
5	B	991/1061 (93%)	871 (88%)	120 (12%)	5 23
6	C	234/274 (85%)	214 (92%)	20 (8%)	10 38
7	E	195/197 (99%)	185 (95%)	10 (5%)	24 57
8	F	76/137 (56%)	70 (92%)	6 (8%)	12 41
9	H	118/128 (92%)	105 (89%)	13 (11%)	6 27

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
10	I	113/116 (97%)	101 (89%)	12 (11%)	6 29
11	J	60/65 (92%)	53 (88%)	7 (12%)	5 24
12	K	99/102 (97%)	86 (87%)	13 (13%)	4 20
13	L	40/57 (70%)	32 (80%)	8 (20%)	1 5
All	All	3159/3657 (86%)	2838 (90%)	321 (10%)	7 30

5 of 321 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
5	B	1220	ARG
10	I	30	ARG
6	C	66	ARG
7	E	165	LEU
12	K	1	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 69 such sidechains are listed below:

Mol	Chain	Res	Type
6	C	112	ASN
7	E	61	GLN
10	I	83	ASN
4	A	1270	ASN
4	A	1265	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	9/10 (90%)	5 (55%)	1 (11%)

All (5) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	2	U
1	R	3	C
1	R	7	A
1	R	8	G
1	R	9	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	R	1	A

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 11 ligands modelled in this entry, 9 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	DGT	B	1308[B]	-	26,33,33	0.79	1 (3%)	32,52,52	1.18	2 (6%)
16	DGT	B	1308[A]	-	26,33,33	0.86	0	32,52,52	1.29	6 (18%)

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	DGT	B	1308[B]	-	-	6/18/34/34	0/3/3/3
16	DGT	B	1308[A]	-	-	5/18/34/34	0/3/3/3

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
16	B	1308[B]	DGT	C6-N1	-2.16	1.34	1.37

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
16	B	1308[A]	DGT	C2'-C1'-N9	-2.48	108.56	114.27
16	B	1308[A]	DGT	C5-C6-N1	2.45	118.28	113.95
16	B	1308[A]	DGT	C2'-C3'-C4'	2.43	107.83	102.76
16	B	1308[B]	DGT	C5-C6-N1	2.41	118.21	113.95
16	B	1308[A]	DGT	O6-C6-C5	-2.40	119.68	124.37

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
16	B	1308[A]	DGT	C5'-O5'-PA-O2A
16	B	1308[A]	DGT	C4'-C5'-O5'-PA
16	B	1308[B]	DGT	C5'-O5'-PA-O3A
16	B	1308[B]	DGT	C3'-C4'-C5'-O5'
16	B	1308[A]	DGT	O4'-C4'-C5'-O5'

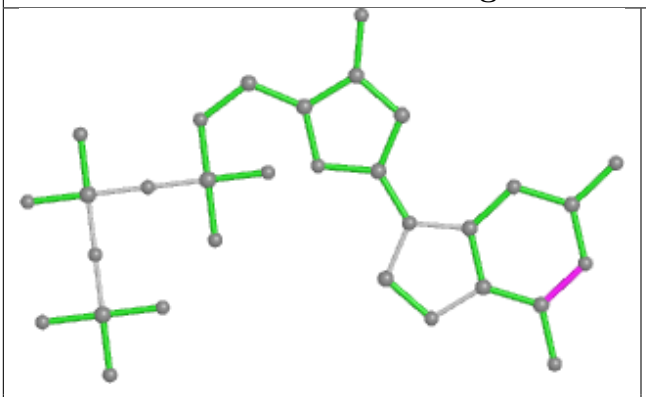
There are no ring outliers.

1 monomer is involved in 2 short contacts:

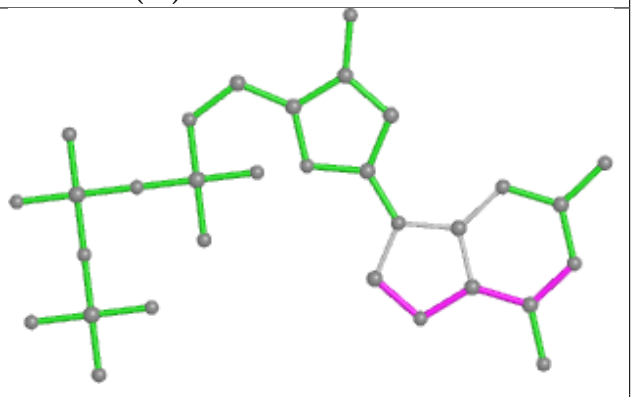
Mol	Chain	Res	Type	Clashes	Symm-Clashes
16	B	1308[A]	DGT	2	0

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.

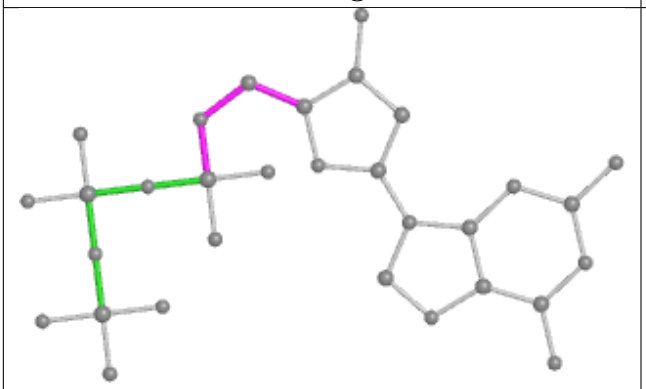
Ligand DGT B 1308 (B)



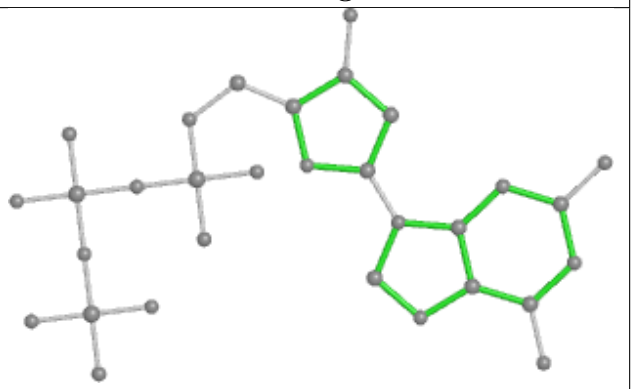
Bond lengths



Bond angles

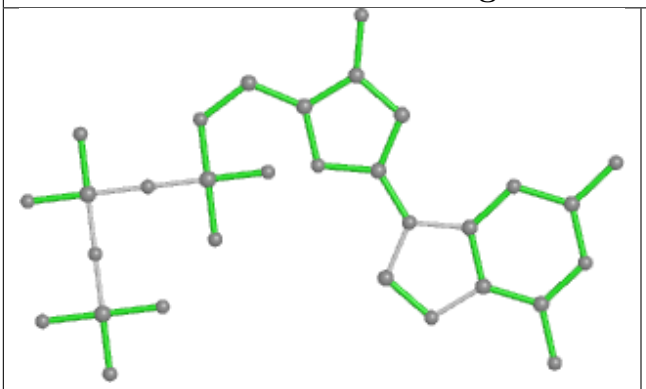


Torsions

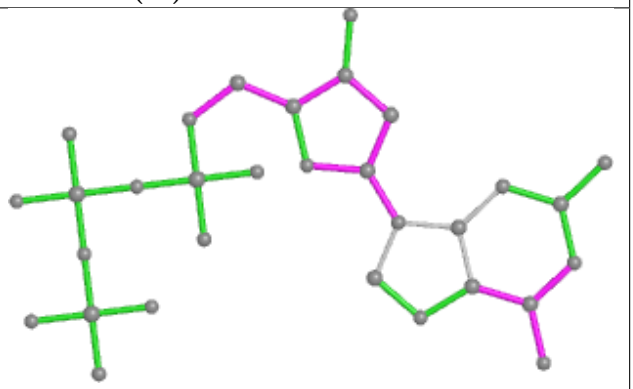


Rings

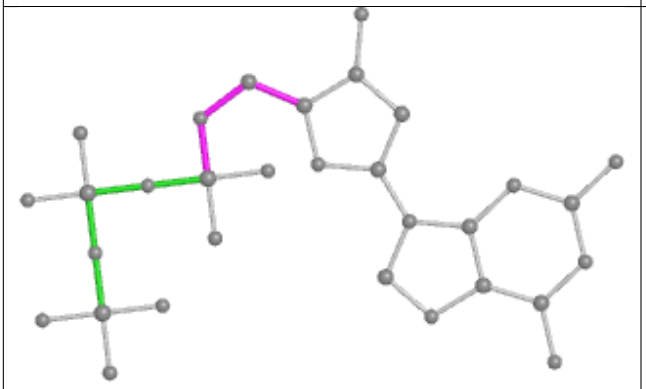
Ligand DGT B 1308 (A)



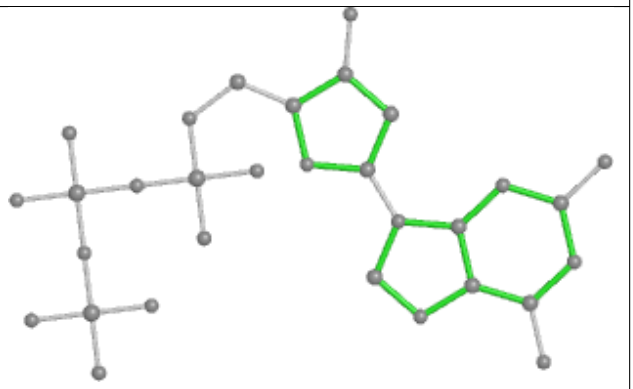
Bond lengths



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	R	10/10 (100%)	0.03	1 (10%) 7 9	72, 93, 202, 234	0
2	T	28/28 (100%)	0.43	1 (3%) 42 42	118, 195, 427, 447	0
3	N	14/14 (100%)	0.44	1 (7%) 16 19	270, 338, 424, 426	0
4	A	1411/1733 (81%)	0.26	71 (5%) 28 29	56, 105, 206, 247	0
5	B	1143/1224 (93%)	0.14	39 (3%) 45 44	27, 88, 156, 185	0
6	C	266/318 (83%)	-0.12	0 100 100	59, 91, 135, 151	0
7	E	213/215 (99%)	0.35	14 (6%) 18 20	96, 142, 228, 233	0
8	F	86/155 (55%)	-0.01	1 (1%) 79 77	83, 117, 137, 141	0
9	H	134/146 (91%)	0.67	12 (8%) 9 12	35, 140, 171, 175	0
10	I	119/122 (97%)	0.04	0 100 100	85, 121, 156, 165	0
11	J	65/70 (92%)	-0.19	0 100 100	60, 82, 110, 117	0
12	K	114/120 (95%)	0.01	0 100 100	77, 102, 119, 122	0
13	L	46/70 (65%)	1.00	6 (13%) 3 5	87, 165, 176, 179	0
All	All	3649/4225 (86%)	0.20	146 (4%) 38 38	27, 103, 198, 447	0

The worst 5 of 146 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
13	L	26	THR	9.2
4	A	69	THR	8.0
4	A	152	VAL	7.6
4	A	144	THR	7.0
4	A	150	THR	6.3

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)

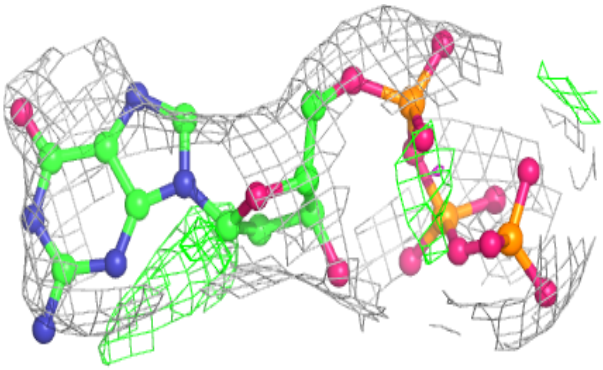
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

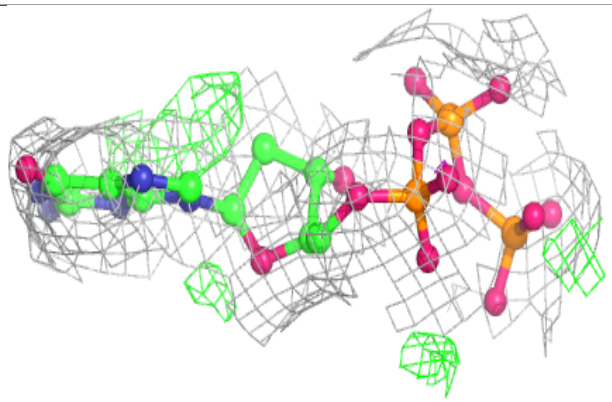
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
16	DGT	B	1308[A]	31/31	0.87	0.41	126,132,152,153	31
16	DGT	B	1308[B]	31/31	0.87	0.41	80,86,113,113	31
14	ZN	A	1734	1/1	0.94	0.05	226,226,226,226	0
14	ZN	A	1735	1/1	0.94	0.10	139,139,139,139	0
14	ZN	B	1307	1/1	0.97	0.13	143,143,143,143	0
14	ZN	I	203	1/1	0.97	0.15	131,131,131,131	0
14	ZN	L	105	1/1	0.98	0.07	190,190,190,190	0
15	MG	A	2001	1/1	0.98	0.38	116,116,116,116	0
14	ZN	I	204	1/1	0.99	0.12	106,106,106,106	0
14	ZN	J	101	1/1	0.99	0.13	105,105,105,105	0
14	ZN	C	319	1/1	0.99	0.10	95,95,95,95	0

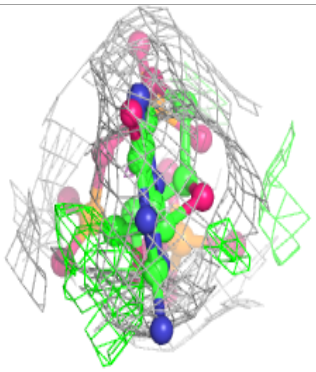
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around DGT B 1308 (A):

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

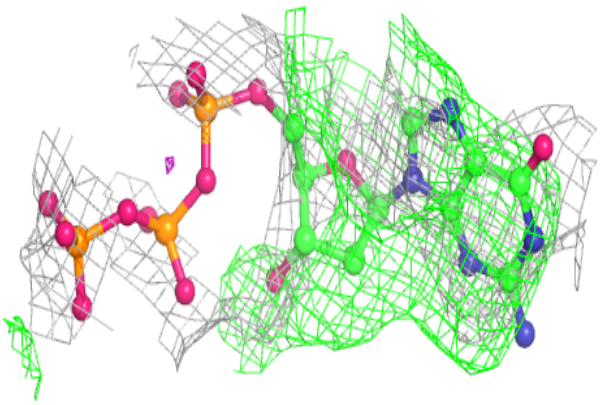


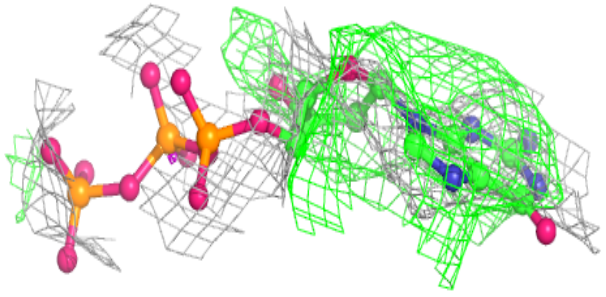


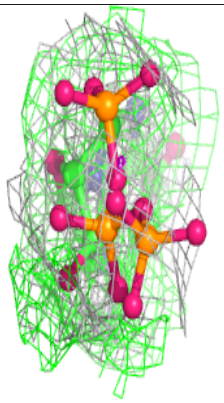


Electron density around DGT B 1308 (B):

$2mF_o$-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.