wwPDB X-ray Structure Validation Summary Report (i)

Nov 16, 2023 – 04:10 AM JST

PDB ID	:	6KQD
Title	:	Thermus thermophilus initial transcription complex comprising sigma A and 5'-OH RNA of 3 nt
Authors	:	Zhang, Y.; Li, L.; Ebright, R.H.
Deposited on	:	2019-08-17
Resolution	:	3.30 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:

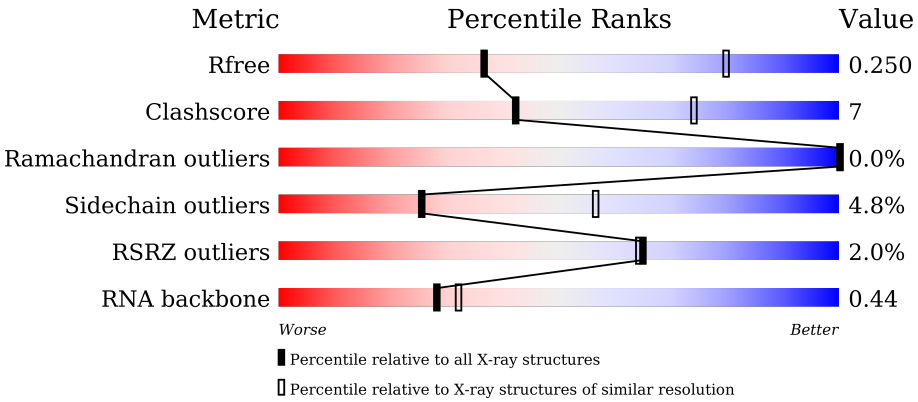
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)
RNA backbone	3102	1117 (3.70-2.90)

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

Mol	Chain	Length	Quality of chain
1	A	315	59% 12% 28%
1	B	315	59% 10% 30%
1	K	315	60% 11% 28%
1	L	315	60% 11% 29%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	1119	2% 81% 17%
2	M	1119	4% 77% 21%
3	D	1524	2% 79% 18%
3	N	1524	% 79% 17%
4	E	99	80% 15% 5%
4	O	99	2% 78% 17% 5%
5	F	443	65% 12% 22%
5	P	443	4% 60% 16% 22%
6	G	21	38% 38% 24%
6	Q	21	48% 29% 24%
7	H	27	33% 56% 11%
7	R	27	44% 44% 11%
8	I	3	33% 33% 33%
8	S	3	33% 33% 33%

2 Entry composition (i)

There are 12 unique types of molecules in this entry. The entry contains 57160 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called DNA-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	226	Total	C	N	O	S	0	0	0
			1782	1138	310	332	2			
1	B	222	Total	C	N	O	S	0	0	0
			1750	1118	304	326	2			
1	K	226	Total	C	N	O	S	0	0	0
			1782	1138	310	332	2			
1	L	225	Total	C	N	O	S	0	0	0
			1773	1133	308	330	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1111	Total	C	N	O	S	0	0	0
			8770	5548	1564	1634	24			
2	M	1111	Total	C	N	O	S	0	0	0
			8761	5542	1564	1631	24			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1486	Total	C	N	O	S	0	0	0
			11738	7441	2067	2195	35			
3	N	1486	Total	C	N	O	S	0	0	0
			11732	7438	2064	2195	35			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	94	Total	C	N	O	S	0	0	0
			761	486	132	139	4			
4	O	94	Total	C	N	O	S	0	0	0
			761	486	132	139	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	346	Total	C	N	O	S	0	0	0
			2804	1769	509	522	4			
5	P	347	Total	C	N	O	S	0	0	0
			2814	1774	510	526	4			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-19	MET	-	initiating methionine	UNP Q5SKW1
F	-18	GLY	-	expression tag	UNP Q5SKW1
F	-17	SER	-	expression tag	UNP Q5SKW1
F	-16	SER	-	expression tag	UNP Q5SKW1
F	-15	HIS	-	expression tag	UNP Q5SKW1
F	-14	HIS	-	expression tag	UNP Q5SKW1
F	-13	HIS	-	expression tag	UNP Q5SKW1
F	-12	HIS	-	expression tag	UNP Q5SKW1
F	-11	HIS	-	expression tag	UNP Q5SKW1
F	-10	HIS	-	expression tag	UNP Q5SKW1
F	-9	SER	-	expression tag	UNP Q5SKW1
F	-8	SER	-	expression tag	UNP Q5SKW1
F	-7	GLY	-	expression tag	UNP Q5SKW1
F	-6	LEU	-	expression tag	UNP Q5SKW1
F	-5	VAL	-	expression tag	UNP Q5SKW1
F	-4	PRO	-	expression tag	UNP Q5SKW1
F	-3	ARG	-	expression tag	UNP Q5SKW1
F	-2	GLY	-	expression tag	UNP Q5SKW1
F	-1	SER	-	expression tag	UNP Q5SKW1
F	0	HIS	-	expression tag	UNP Q5SKW1
P	-19	MET	-	initiating methionine	UNP Q5SKW1
P	-18	GLY	-	expression tag	UNP Q5SKW1
P	-17	SER	-	expression tag	UNP Q5SKW1
P	-16	SER	-	expression tag	UNP Q5SKW1
P	-15	HIS	-	expression tag	UNP Q5SKW1
P	-14	HIS	-	expression tag	UNP Q5SKW1
P	-13	HIS	-	expression tag	UNP Q5SKW1
P	-12	HIS	-	expression tag	UNP Q5SKW1
P	-11	HIS	-	expression tag	UNP Q5SKW1
P	-10	HIS	-	expression tag	UNP Q5SKW1
P	-9	SER	-	expression tag	UNP Q5SKW1
P	-8	SER	-	expression tag	UNP Q5SKW1
P	-7	GLY	-	expression tag	UNP Q5SKW1
P	-6	LEU	-	expression tag	UNP Q5SKW1

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
P	-5	VAL	-	expression tag	UNP Q5SKW1
P	-4	PRO	-	expression tag	UNP Q5SKW1
P	-3	ARG	-	expression tag	UNP Q5SKW1
P	-2	GLY	-	expression tag	UNP Q5SKW1
P	-1	SER	-	expression tag	UNP Q5SKW1
P	0	HIS	-	expression tag	UNP Q5SKW1

- Molecule 6 is a DNA chain called DNA (5'-D(*CP*CP*T*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*CP*CP*AP*GP*GP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	16	Total	C	N	O	P	0	0	0
			325	155	61	94	15			
6	Q	16	Total	C	N	O	P	0	0	0
			325	155	61	94	15			

- Molecule 7 is a DNA chain called DNA (5'-D(*TP*AP*TP*AP*AP*TP*GP*GP*GP*AP*GP*CP*TP*GP*TP*CP*AP*CP*GP*GP*AP*TP*GP*CP*AP*GP*G*)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	24	Total	C	N	O	P	0	0	0
			495	236	94	142	23			
7	R	24	Total	C	N	O	P	0	0	0
			495	236	94	142	23			

- Molecule 8 is a RNA chain called RNA (5'-R(*GP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	I	3	Total	C	N	O	P	0	0	0
			65	30	15	18	2			
8	S	3	Total	C	N	O	P	0	0	0
			65	30	15	18	2			

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	B	1	Total	Mg	0	0
			1	1		
9	D	2	Total	Mg	0	0
			2	2		
9	F	1	Total	Mg	0	0
			1	1		

Continued on next page...

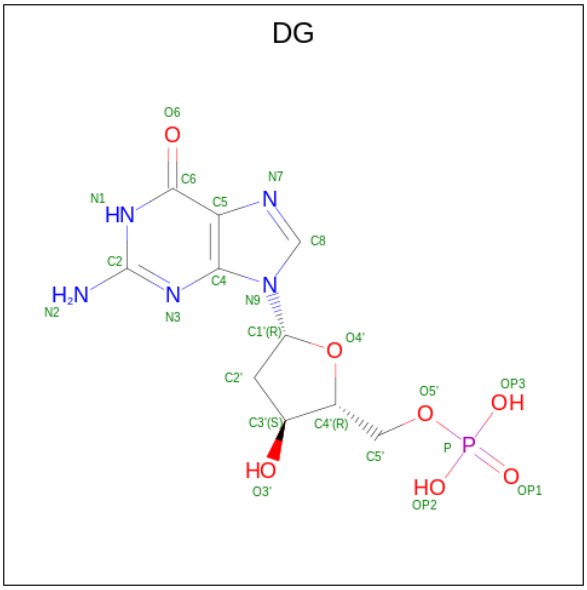
Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	K	1	Total Mg 1 1	0	0
9	N	2	Total Mg 2 2	0	0

- Molecule 10 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 11 is 2'-DEOXYGUANOSINE-5'-MONOPHOSPHATE (three-letter code: DG) (formula: $C_{10}H_{14}N_5O_7P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
11	G	1	Total 22	C 10	N 5	O 6	P 1	0	0
11	Q	1	Total 22	C 10	N 5	O 6	P 1	0	0

- Molecule 12 is water.

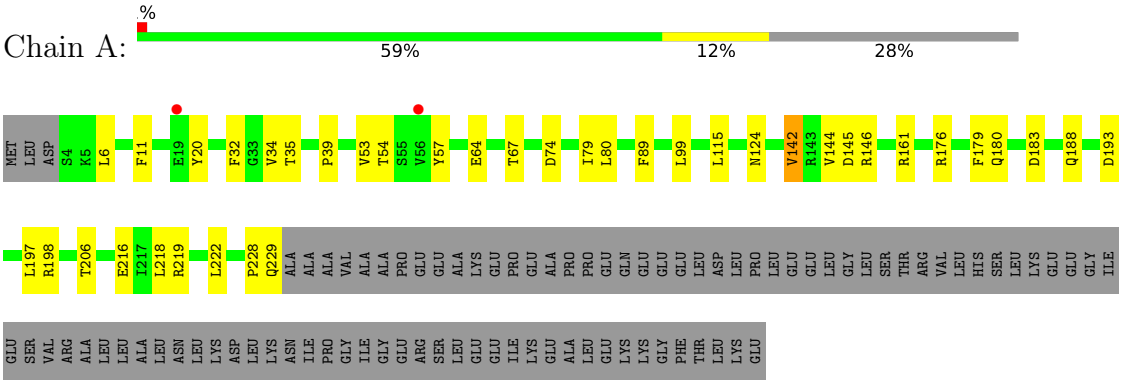
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
12	A	5	Total O 5 5	0	0
12	B	2	Total O 2 2	0	0
12	C	23	Total O 23 23	0	0
12	D	26	Total O 26 26	0	0
12	E	1	Total O 1 1	0	0
12	F	4	Total O 4 4	0	0
12	G	2	Total O 2 2	0	0
12	K	6	Total O 6 6	0	0
12	L	4	Total O 4 4	0	0
12	M	14	Total O 14 14	0	0
12	N	14	Total O 14 14	0	0
12	O	1	Total O 1 1	0	0
12	P	4	Total O 4 4	0	0
12	Q	1	Total O 1 1	0	0

3 Residue-property plots (i)

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

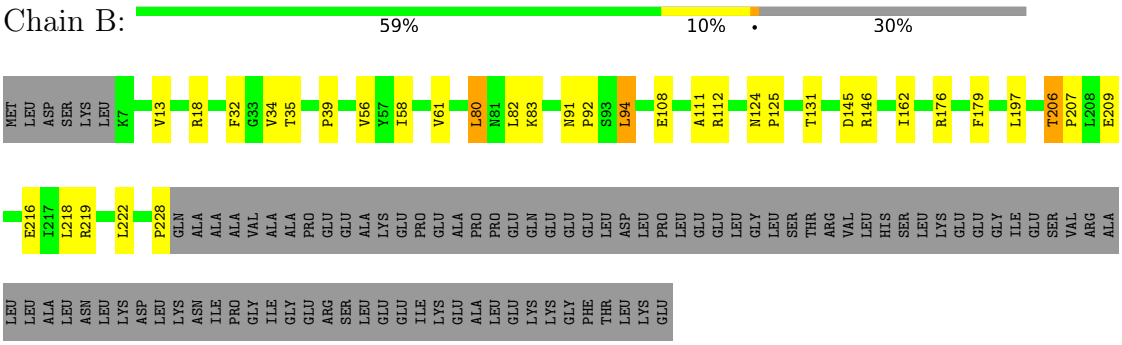
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain A:



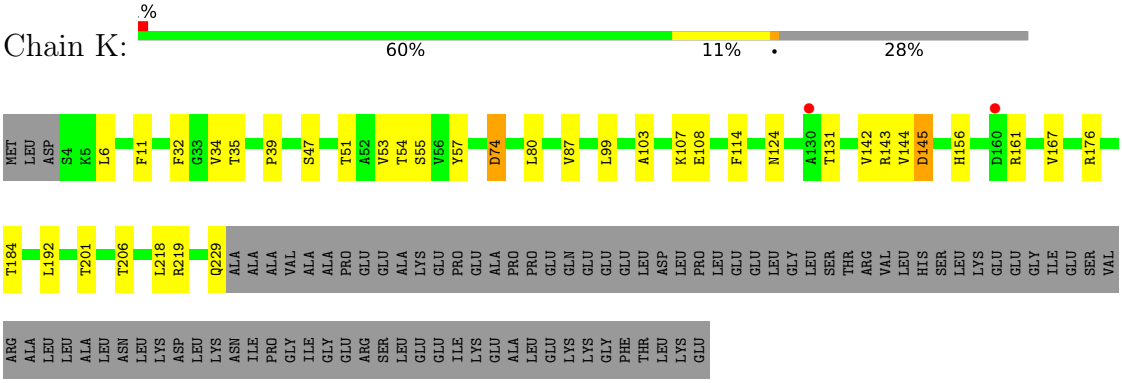
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain B:



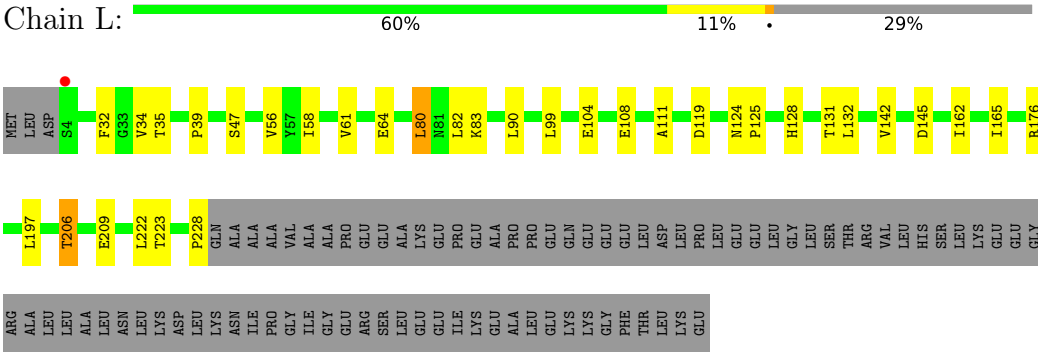
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain K:



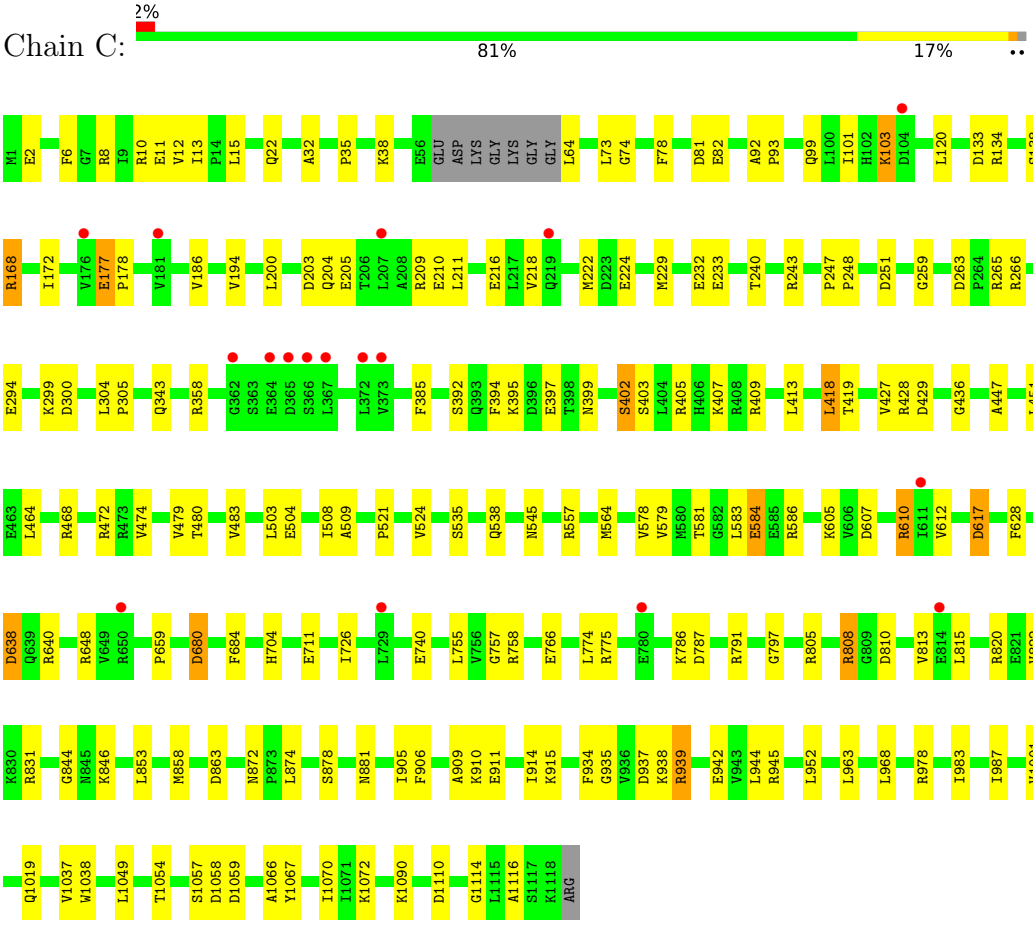
• Molecule 1: DNA-directed RNA polymerase subunit alpha

Chain L:



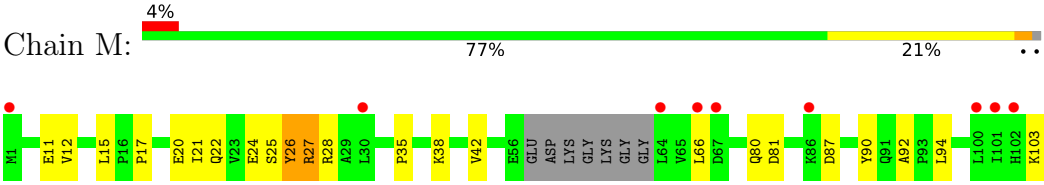
• Molecule 2: DNA-directed RNA polymerase subunit beta

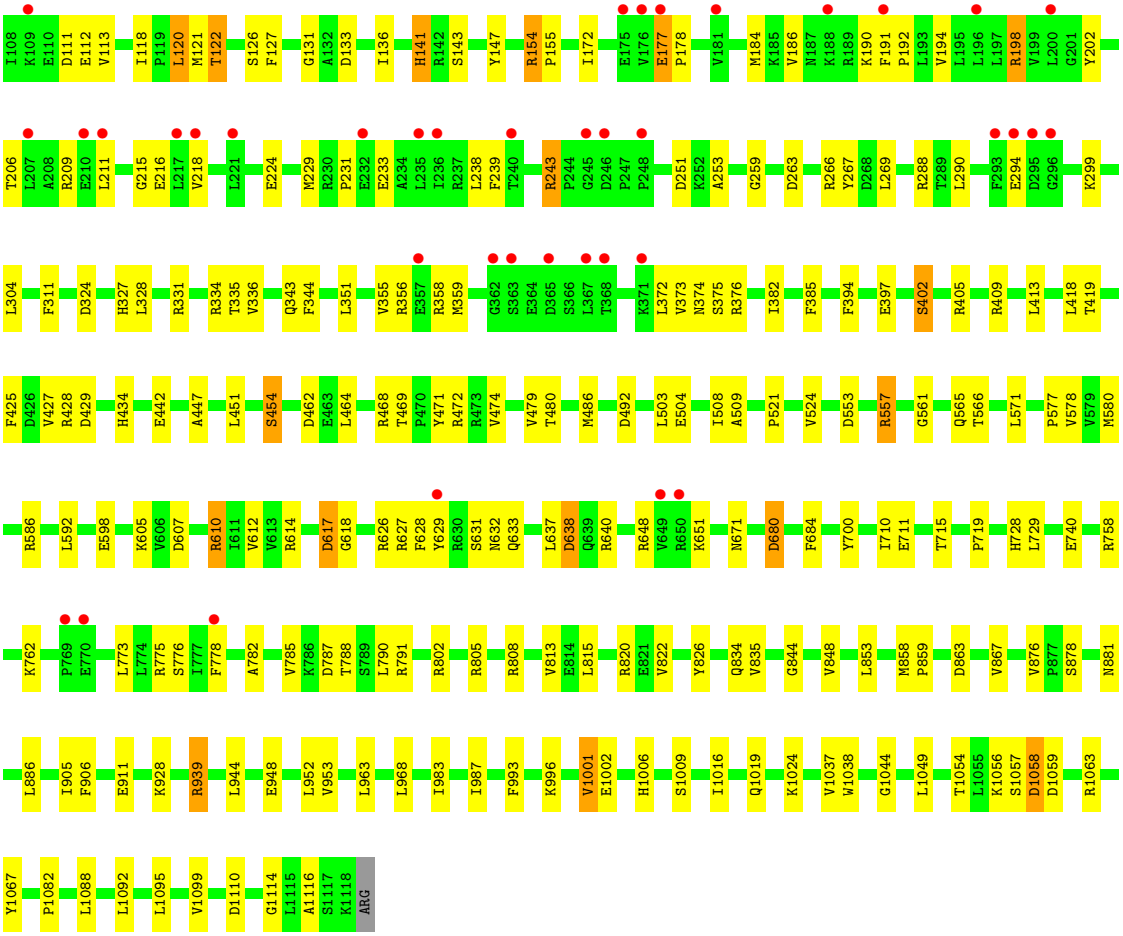
Chain C:



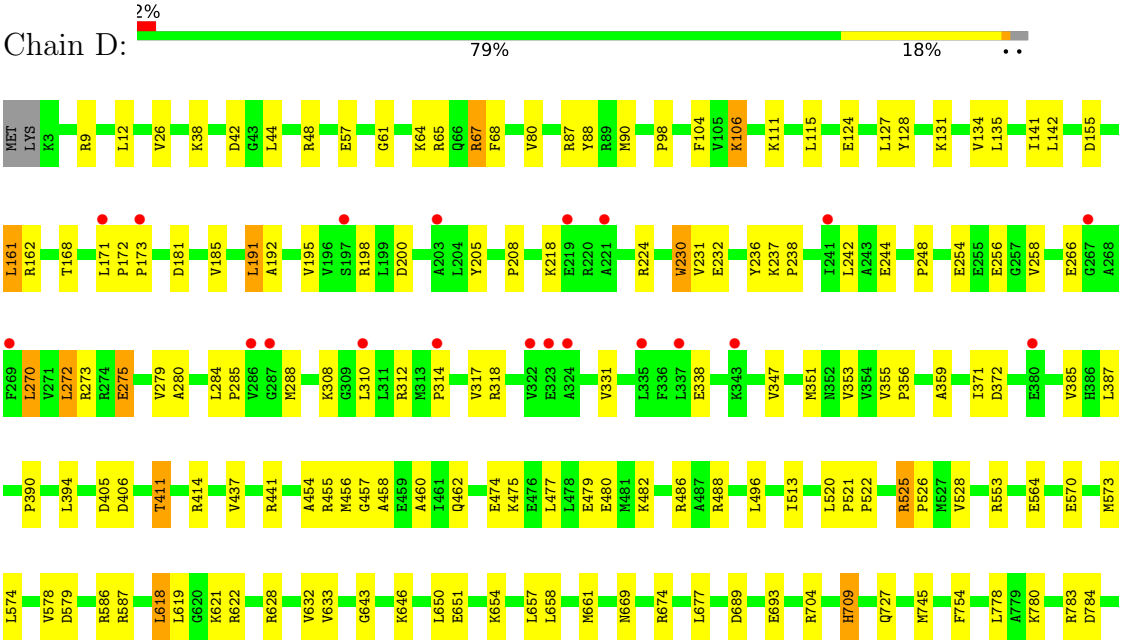
• Molecule 2: DNA-directed RNA polymerase subunit beta

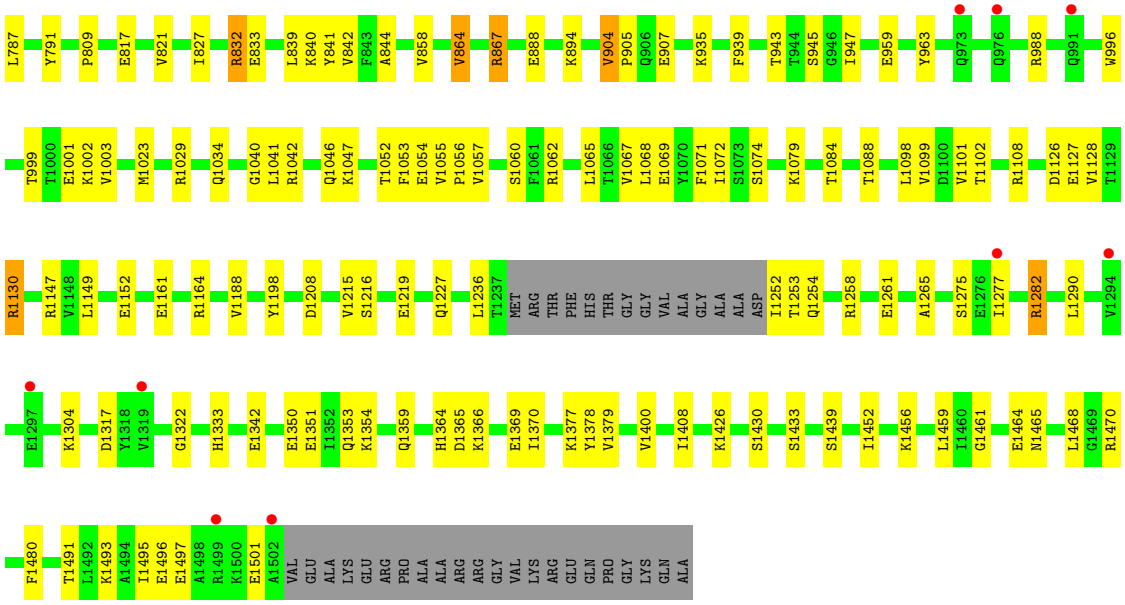
Chain M:



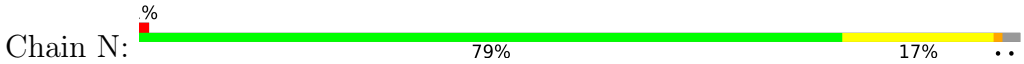


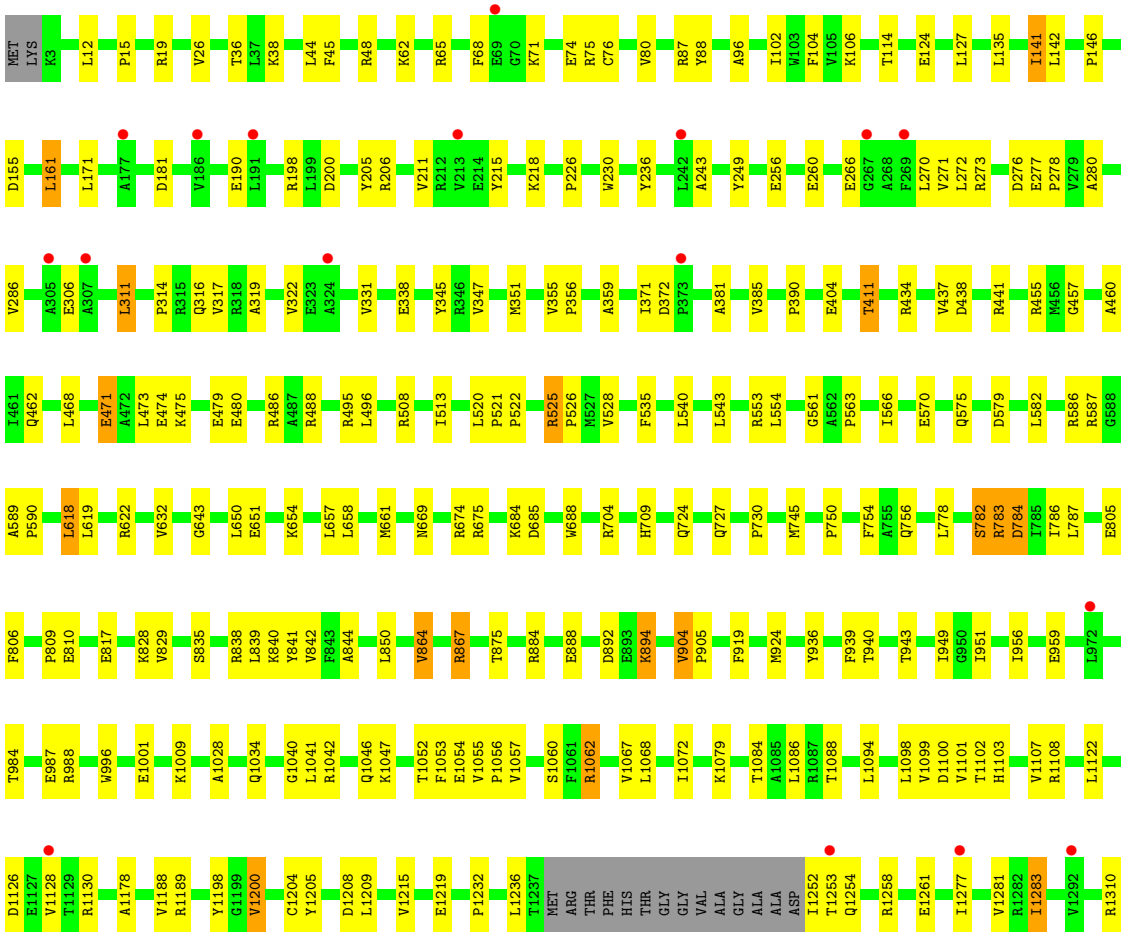
• Molecule 3: DNA-directed RNA polymerase subunit beta'

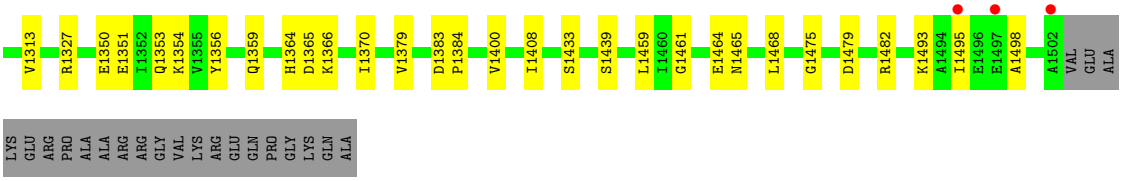




• Molecule 3: DNA-directed RNA polymerase subunit beta'





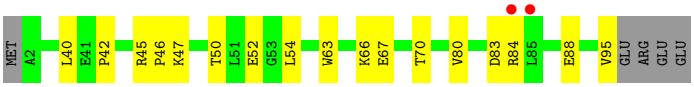


• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E:

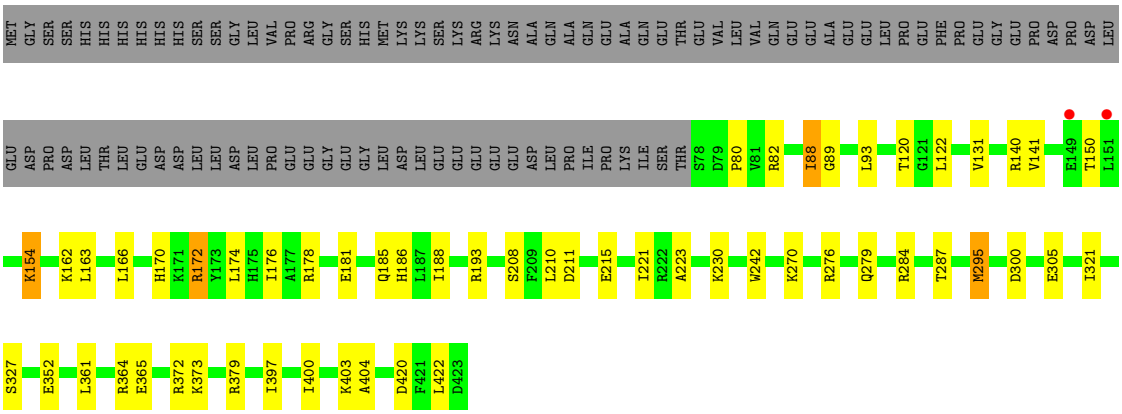
• Molecule 4: DNA-directed RNA polymerase subunit omega

Chain O:



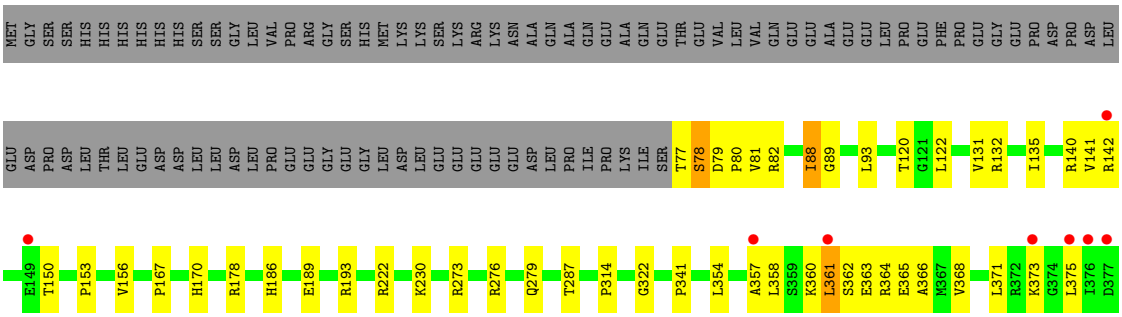
• Molecule 5: RNA polymerase sigma factor SigA

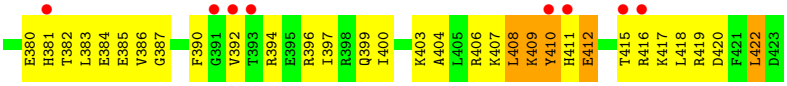
Chain F:



• Molecule 5: RNA polymerase sigma factor SigA

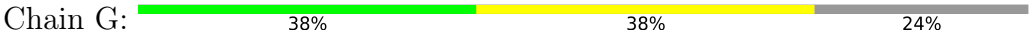
Chain P:





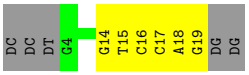
• Molecule 6: DNA (5'-D(*CP*CP*T*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*CP*CP*AP*GP*GP*G)-3')

Chain G:



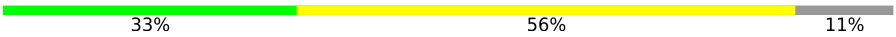
• Molecule 6: DNA (5'-D(*CP*CP*T*GP*CP*AP*TP*CP*CP*GP*TP*GP*AP*GP*TP*CP*CP*AP*GP*GP*G)-3')

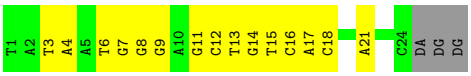
Chain Q:



• Molecule 7: DNA (5'-D(*TP*AP*TP*AP*AP*TP*GP*GP*GP*AP*GP*CP*TP*GP*TP*CP*AP*CP*GP*GP*AP*TP*GP*CP*AP*GP*G*)-3')

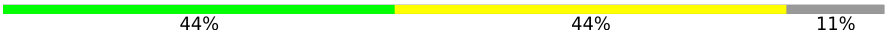
Chain H:

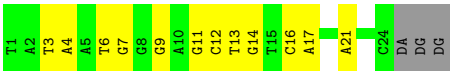




• Molecule 7: DNA (5'-D(*TP*AP*TP*AP*AP*TP*GP*GP*GP*AP*GP*CP*TP*GP*TP*CP*AP*CP*GP*GP*AP*TP*GP*CP*AP*GP*G*)-3')

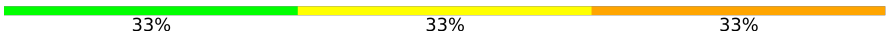
Chain R:





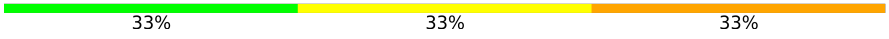
• Molecule 8: RNA (5'-R(*GP*GP*A)-3')

Chain I:



• Molecule 8: RNA (5'-R(*GP*GP*A)-3')

Chain S:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	185.50Å 104.04Å 297.44Å 90.00° 98.50° 90.00°	Depositor
Resolution (Å)	44.35 – 3.30 44.35 – 3.30	Depositor EDS
% Data completeness (in resolution range)	91.3 (44.35-3.30) 91.5 (44.35-3.30)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.29 (at 3.32Å)	Xtriage
Refinement program	PHENIX 1.14_3260	Depositor
R, R_{free}	0.207 , 0.250 0.207 , 0.250	Depositor DCC
R_{free} test set	1338 reflections (0.86%)	wwPDB-VP
Wilson B-factor (Å^2)	77.2	Xtriage
Anisotropy	0.549	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 49.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	57160	wwPDB-VP
Average B, all atoms (Å^2)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 44.47 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.5182e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: MG, ZN

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	0.27	0/1814	0.46	0/2466
1	B	0.26	0/1782	0.46	0/2424
1	K	0.27	0/1814	0.46	0/2466
1	L	0.26	0/1805	0.46	0/2454
2	C	0.27	0/8937	0.45	0/12087
2	M	0.28	0/8927	0.46	0/12073
3	D	0.28	0/11944	0.45	0/16149
3	N	0.26	0/11938	0.44	0/16142
4	E	0.26	0/775	0.41	0/1045
4	O	0.25	0/775	0.41	0/1045
5	F	0.25	0/2849	0.41	0/3833
5	P	0.26	0/2859	0.42	0/3847
6	G	0.69	0/364	0.90	0/560
6	Q	0.58	0/364	0.90	0/560
7	H	0.67	0/556	0.98	0/858
7	R	0.64	0/556	0.97	0/858
8	I	0.36	0/73	0.85	0/113
8	S	0.42	0/73	1.14	0/113
All	All	0.29	0/58205	0.48	0/79093

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	1782	0	1834	21	0
1	B	1750	0	1797	23	0
1	K	1782	0	1834	19	0
1	L	1773	0	1826	26	0
2	C	8770	0	8874	114	0
2	M	8761	0	8863	174	0
3	D	11738	0	11971	167	0
3	N	11732	0	11960	181	0
4	E	761	0	778	10	0
4	O	761	0	778	11	0
5	F	2804	0	2880	36	0
5	P	2814	0	2889	60	0
6	G	325	0	181	9	0
6	Q	325	0	181	6	0
7	H	495	0	272	18	0
7	R	495	0	272	16	0
8	I	65	0	34	2	0
8	S	65	0	34	2	0
9	B	1	0	0	0	0
9	D	2	0	0	0	0
9	F	1	0	0	0	0
9	K	1	0	0	0	0
9	N	2	0	0	0	0
10	D	2	0	0	0	0
10	N	2	0	0	0	0
11	G	22	0	12	3	0
11	Q	22	0	12	1	0
12	A	5	0	0	0	0
12	B	2	0	0	0	0
12	C	23	0	0	0	0
12	D	26	0	0	0	0
12	E	1	0	0	0	0
12	F	4	0	0	0	0
12	G	2	0	0	0	0
12	K	6	0	0	0	0
12	L	4	0	0	0	0
12	M	14	0	0	1	0
12	N	14	0	0	1	0
12	O	1	0	0	0	0
12	P	4	0	0	0	0
12	Q	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	57160	0	57282	784	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 7.

The worst 5 of 784 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:M:90:TYR:CE2	2:M:120:LEU:HD12	1.75	1.21
3:N:243:ALA:HB3	3:N:311:LEU:HD21	1.27	1.11
2:M:90:TYR:CE2	2:M:120:LEU:CD1	2.33	1.11
2:M:90:TYR:CD2	2:M:120:LEU:HD12	1.97	0.99
2:M:27:ARG:HB3	2:M:27:ARG:NH1	1.78	0.96

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	224/315 (71%)	220 (98%)	4 (2%)	0	100	100
1	B	220/315 (70%)	216 (98%)	4 (2%)	0	100	100
1	K	224/315 (71%)	222 (99%)	2 (1%)	0	100	100
1	L	223/315 (71%)	218 (98%)	5 (2%)	0	100	100
2	C	1107/1119 (99%)	1089 (98%)	18 (2%)	0	100	100
2	M	1107/1119 (99%)	1084 (98%)	23 (2%)	0	100	100
3	D	1482/1524 (97%)	1462 (99%)	20 (1%)	0	100	100
3	N	1482/1524 (97%)	1463 (99%)	19 (1%)	0	100	100
4	E	92/99 (93%)	91 (99%)	1 (1%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
4	O	92/99 (93%)	91 (99%)	1 (1%)	0	100	100
5	F	344/443 (78%)	343 (100%)	1 (0%)	0	100	100
5	P	345/443 (78%)	341 (99%)	3 (1%)	1 (0%)	41	71
All	All	6942/7630 (91%)	6840 (98%)	101 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
5	P	80	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	
1	A	199/273 (73%)	189 (95%)	10 (5%)	24	55
1	B	195/273 (71%)	189 (97%)	6 (3%)	40	67
1	K	199/273 (73%)	189 (95%)	10 (5%)	24	55
1	L	198/273 (72%)	192 (97%)	6 (3%)	41	68
2	C	936/941 (100%)	889 (95%)	47 (5%)	24	55
2	M	934/941 (99%)	884 (95%)	50 (5%)	22	53
3	D	1253/1279 (98%)	1191 (95%)	62 (5%)	25	56
3	N	1252/1279 (98%)	1192 (95%)	60 (5%)	25	56
4	E	83/88 (94%)	81 (98%)	2 (2%)	49	73
4	O	83/88 (94%)	81 (98%)	2 (2%)	49	73
5	F	300/388 (77%)	288 (96%)	12 (4%)	31	61
5	P	302/388 (78%)	284 (94%)	18 (6%)	19	49
All	All	5934/6484 (92%)	5649 (95%)	285 (5%)	25	56

5 of 285 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	N	525	ARG
3	N	754	PHE
3	N	1277	ILE
3	D	817	GLU
3	D	754	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
3	D	1333	HIS
4	E	33	HIS
2	M	22	GLN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	I	2/3 (66%)	1 (50%)	0
8	S	2/3 (66%)	1 (50%)	0
All	All	4/6 (66%)	2 (50%)	0

All (2) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
8	I	3	A
8	S	3	A

There are no RNA pucker outliers to report.

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 13 ligands modelled in this entry, 11 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	$\#\|Z\| > 2$
11	DG	G	101	-	18,24,25	0.93	1 (5%)	19,35,38	0.79	1 (5%)
11	DG	Q	101	-	18,24,25	1.02	2 (11%)	19,35,38	0.69	1 (5%)

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
11	DG	G	101	-	-	2/3/21/22	0/3/3/3
11	DG	Q	101	-	-	0/3/21/22	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	Q	101	DG	C5-C6	-2.49	1.42	1.47
11	G	101	DG	C5-C6	-2.20	1.42	1.47
11	Q	101	DG	C8-N7	-2.16	1.31	1.35

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
11	G	101	DG	O6-C6-C5	2.26	128.78	124.37
11	Q	101	DG	O6-C6-C5	2.04	128.36	124.37

There are no chirality outliers.

All (2) torsion outliers are listed below:

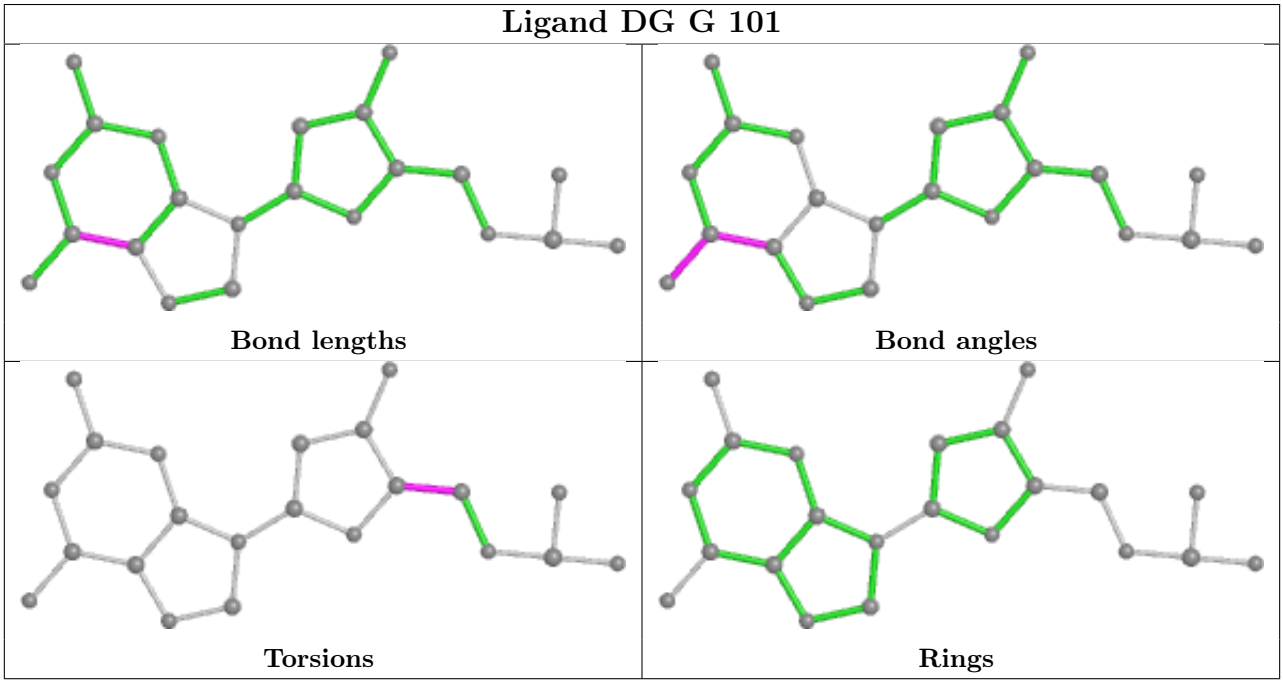
Mol	Chain	Res	Type	Atoms
11	G	101	DG	O4'-C4'-C5'-O5'
11	G	101	DG	C3'-C4'-C5'-O5'

There are no ring outliers.

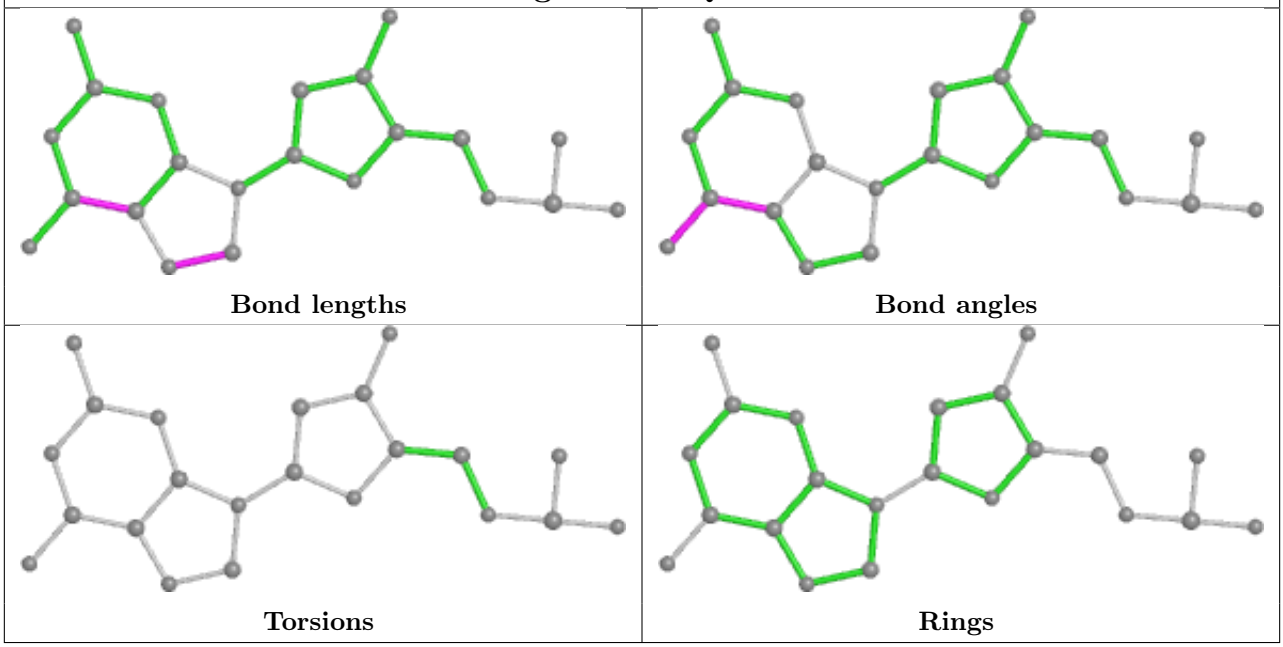
2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	G	101	DG	3	0
11	Q	101	DG	1	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 DG Q 101



5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	226/315 (71%)	-0.25	2 (0%) 84 84	47, 73, 98, 113	0
1	B	222/315 (70%)	-0.19	0 100 100	47, 81, 116, 144	0
1	K	226/315 (71%)	-0.13	2 (0%) 84 84	54, 84, 109, 120	0
1	L	225/315 (71%)	-0.19	1 (0%) 92 93	50, 89, 124, 145	0
2	C	1111/1119 (99%)	-0.14	17 (1%) 73 72	29, 65, 124, 151	0
2	M	1111/1119 (99%)	0.08	50 (4%) 33 32	34, 88, 154, 168	0
3	D	1486/1524 (97%)	-0.10	29 (1%) 65 64	27, 68, 131, 177	1 (0%)
3	N	1486/1524 (97%)	-0.07	20 (1%) 77 77	34, 75, 131, 182	1 (0%)
4	E	94/99 (94%)	-0.20	0 100 100	43, 69, 114, 120	0
4	O	94/99 (94%)	-0.25	2 (2%) 63 62	54, 83, 126, 140	0
5	F	346/443 (78%)	-0.16	2 (0%) 89 90	43, 78, 129, 142	0
5	P	347/443 (78%)	0.08	16 (4%) 32 30	56, 94, 163, 181	0
6	G	16/21 (76%)	-0.49	0 100 100	44, 84, 174, 175	0
6	Q	16/21 (76%)	-0.60	0 100 100	54, 94, 182, 184	0
7	H	24/27 (88%)	-0.50	0 100 100	71, 109, 162, 194	0
7	R	24/27 (88%)	-0.37	0 100 100	72, 119, 172, 204	0
8	I	3/3 (100%)	-0.55	0 100 100	57, 57, 60, 69	0
8	S	3/3 (100%)	-0.49	0 100 100	66, 66, 69, 78	0
All	All	7060/7732 (91%)	-0.08	141 (1%) 65 64	27, 78, 137, 204	2 (0%)

The worst 5 of 141 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	P	375	LEU	5.5
3	D	322	VAL	5.0
5	P	377	ASP	5.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	N	191	LEU	4.9
2	M	64	LEU	4.8

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)

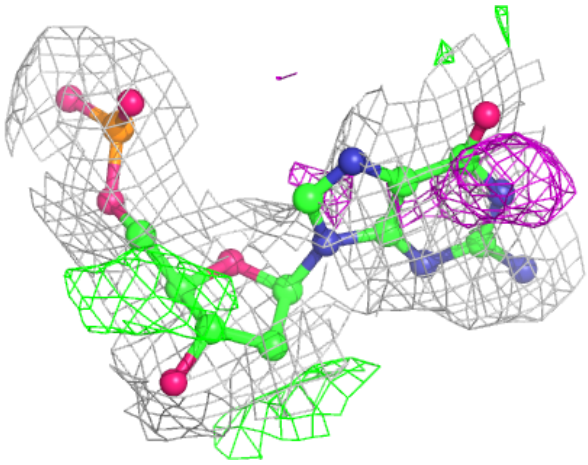
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

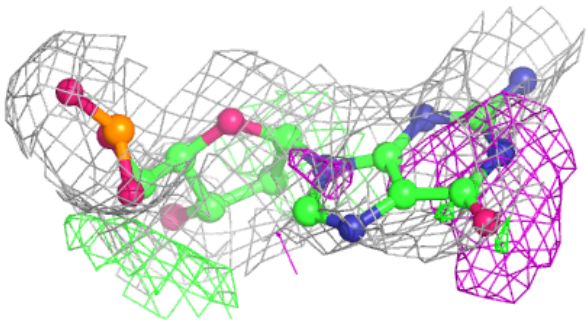
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
11	DG	Q	101	22/23	0.60	0.31	89,113,119,120	0
9	MG	B	2001	1/1	0.66	0.33	80,80,80,80	0
11	DG	G	101	22/23	0.77	0.25	71,100,108,124	0
9	MG	F	2001	1/1	0.85	0.10	64,64,64,64	0
9	MG	K	1001	1/1	0.93	0.27	76,76,76,76	0
10	ZN	N	2002	1/1	0.96	0.05	120,120,120,120	0
9	MG	N	2003	1/1	0.97	0.25	44,44,44,44	0
9	MG	N	2004	1/1	0.98	0.31	61,61,61,61	0
9	MG	D	2004	1/1	0.98	0.48	42,42,42,42	0
10	ZN	N	2001	1/1	0.99	0.20	55,55,55,55	0
9	MG	D	2003	1/1	0.99	0.22	36,36,36,36	0
10	ZN	D	2001	1/1	0.99	0.18	49,49,49,49	0
10	ZN	D	2002	1/1	0.99	0.07	88,88,88,88	0

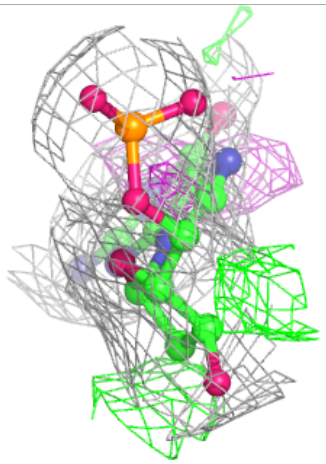
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around DG Q 101:

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

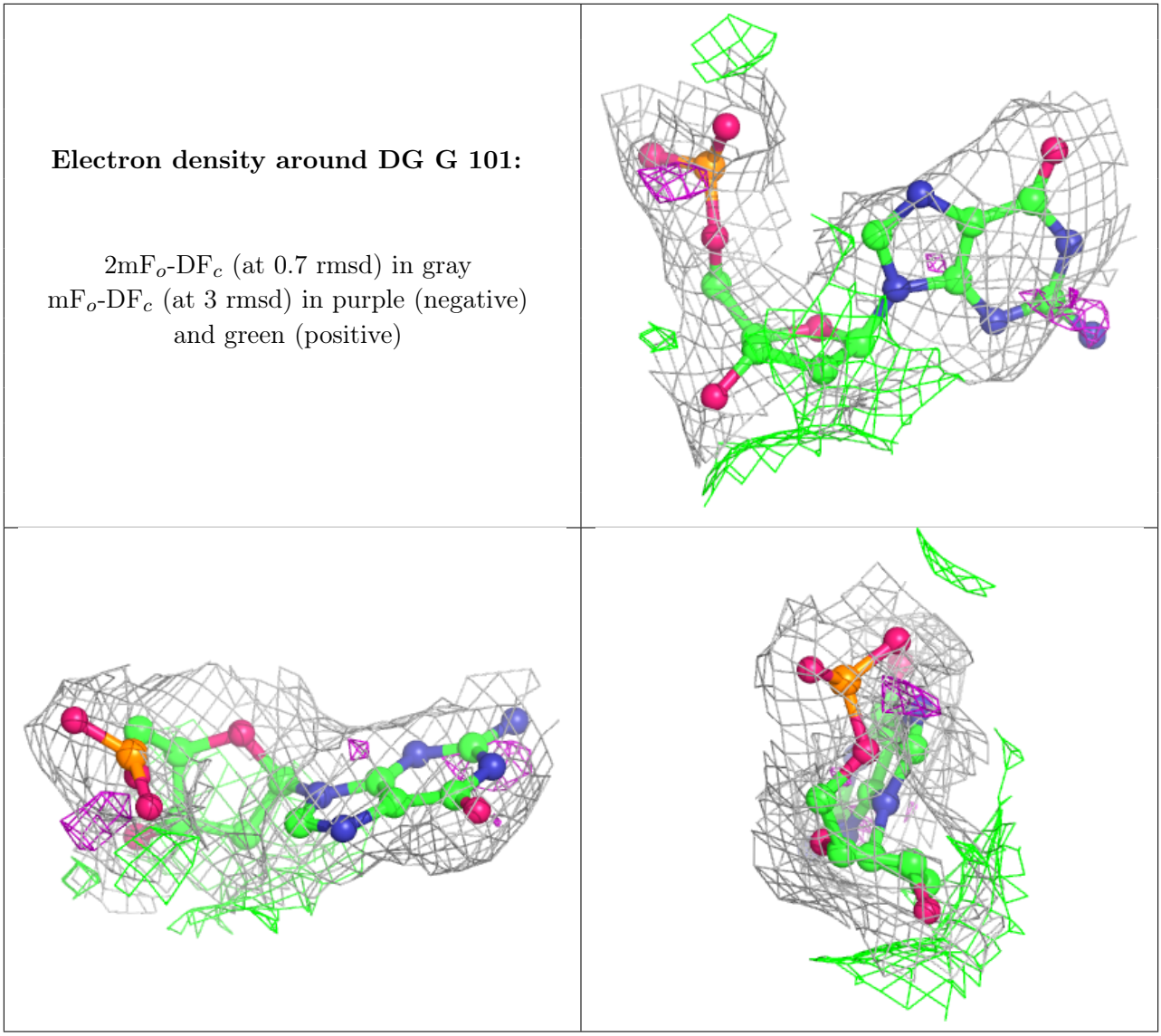






Electron density around DG G 101:

$2mF_o$-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

There are no such residues in this entry.