wwPDB X-ray Structure Validation Summary Report (i)

Feb 3, 2024 – 02:31 PM EST

PDB ID	:	1L9Z
Title	:	Thermus aquaticus RNA Polymerase Holoenzyme/Fork-Junction Promoter DNA Complex at 6.5 A Resolution
Authors	:	Murakami, K.S.; Masuda, S.; Campbell, E.A.; Muzzin, O.; Darst, S.A.
Deposited on	:	2002-03-27
Resolution	:	6.50 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:

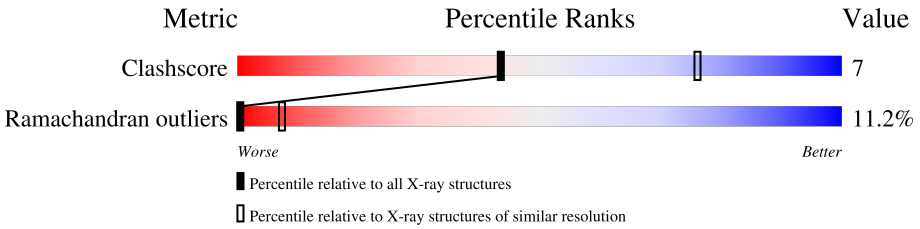
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 6.50 Å.

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



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	1064 (9.00-3.90)
Ramachandran outliers	138981	1012 (9.00-3.88)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	U	35	26% 54% 20%
2	T	30	60% 13% 27%
3	A	314	62% 9% 29%
3	B	314	61% 9% 30%
4	C	1118	86% 11% .
5	D	1524	66% 11% 22%
6	E	99	88% 5% 7%
7	H	438	63% 10% 27%

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 10078 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 DNA chain called nontemplate DNA strand.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
1	U	35	Total	C	O	P	0	0	0
			382	175	173	34			

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

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	T	30	Total	C	O	P	0	0	0
			327	150	148	29			

- Molecule 3 is a protein called RNA POLYMERASE, ALPHA SUBUNIT.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
3	A	224	Total	C	N	0	0	0
			672	448	224			
3	B	220	Total	C	N	0	0	0
			660	440	220			

- Molecule 4 is a protein called RNA POLYMERASE, BETA SUBUNIT.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
4	C	1084	Total	C	N	0	0	0
			3252	2168	1084			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	?	-	GLU	deletion	UNP Q9KWU7

- Molecule 5 is a protein called RNA POLYMERASE, BETA-PRIME SUBUNIT.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
5	D	1183	Total 3549	C 2366	N 1183	0	0	0

- Molecule 6 is a protein called RNA POLYMERASE, OMEGA SUBUNIT.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
6	E	92	Total 276	C 184	N 92	0	0	0

- Molecule 7 is a protein called SIGMA FACTOR SIGA.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace
7	H	319	Total 957	C 638	N 319	0	0	0

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

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

- Molecule 9 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

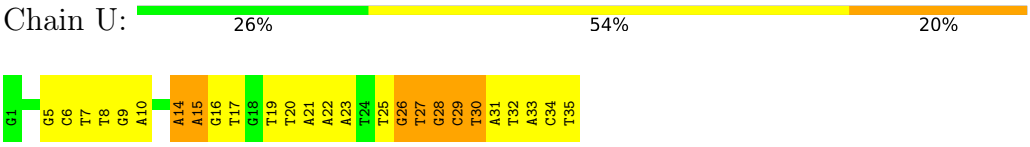
3 Residue-property plots (i)

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. 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. 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.

Note EDS was not executed.

• Molecule 1: nontemplate DNA strand

Chain U:

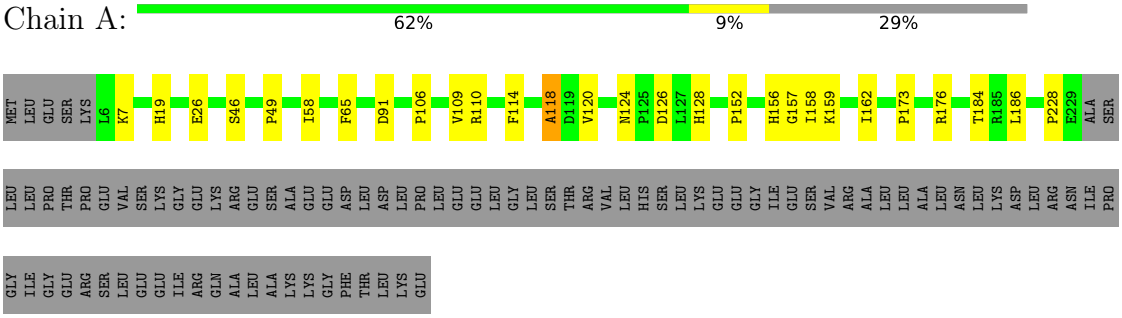


• Molecule 2: template DNA strand

Chain T:

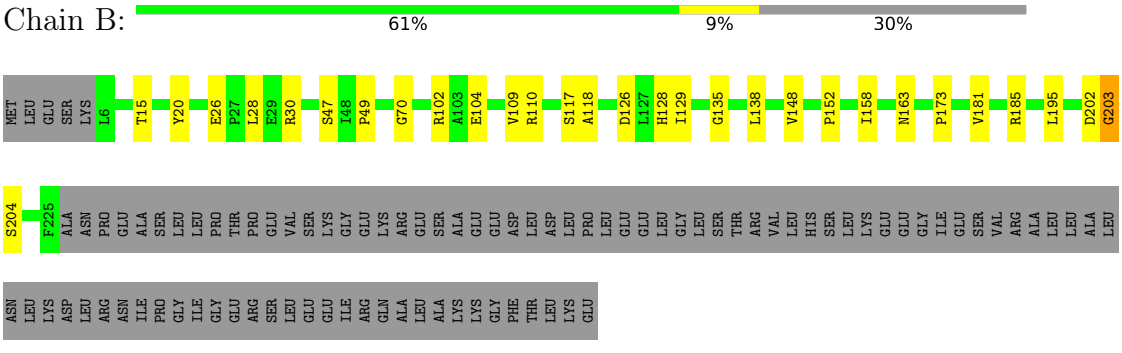
• Molecule 3: RNA POLYMERASE, ALPHA SUBUNIT

Chain A:



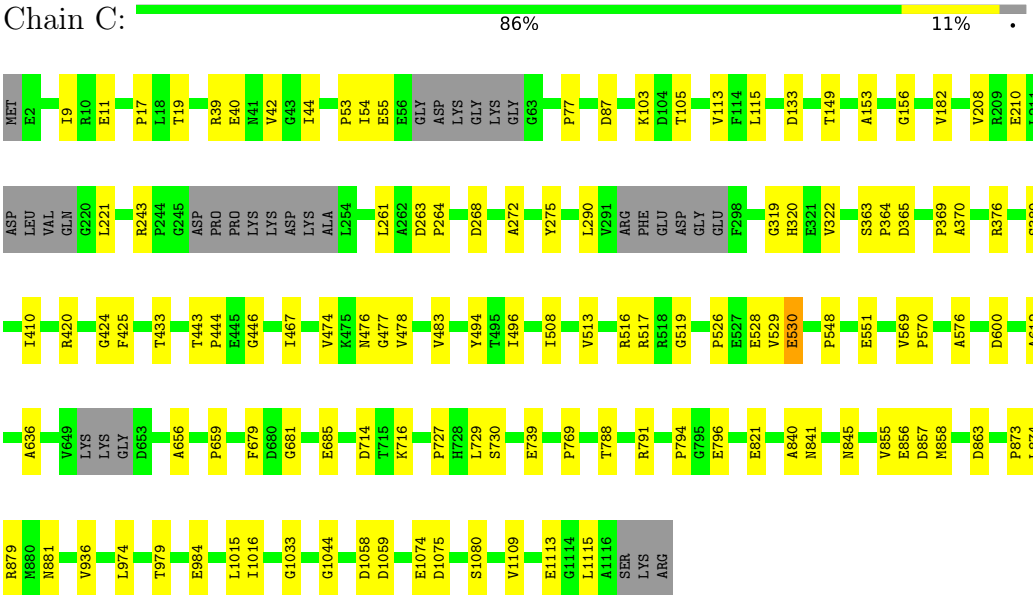
• Molecule 3: RNA POLYMERASE, ALPHA SUBUNIT

Chain B:



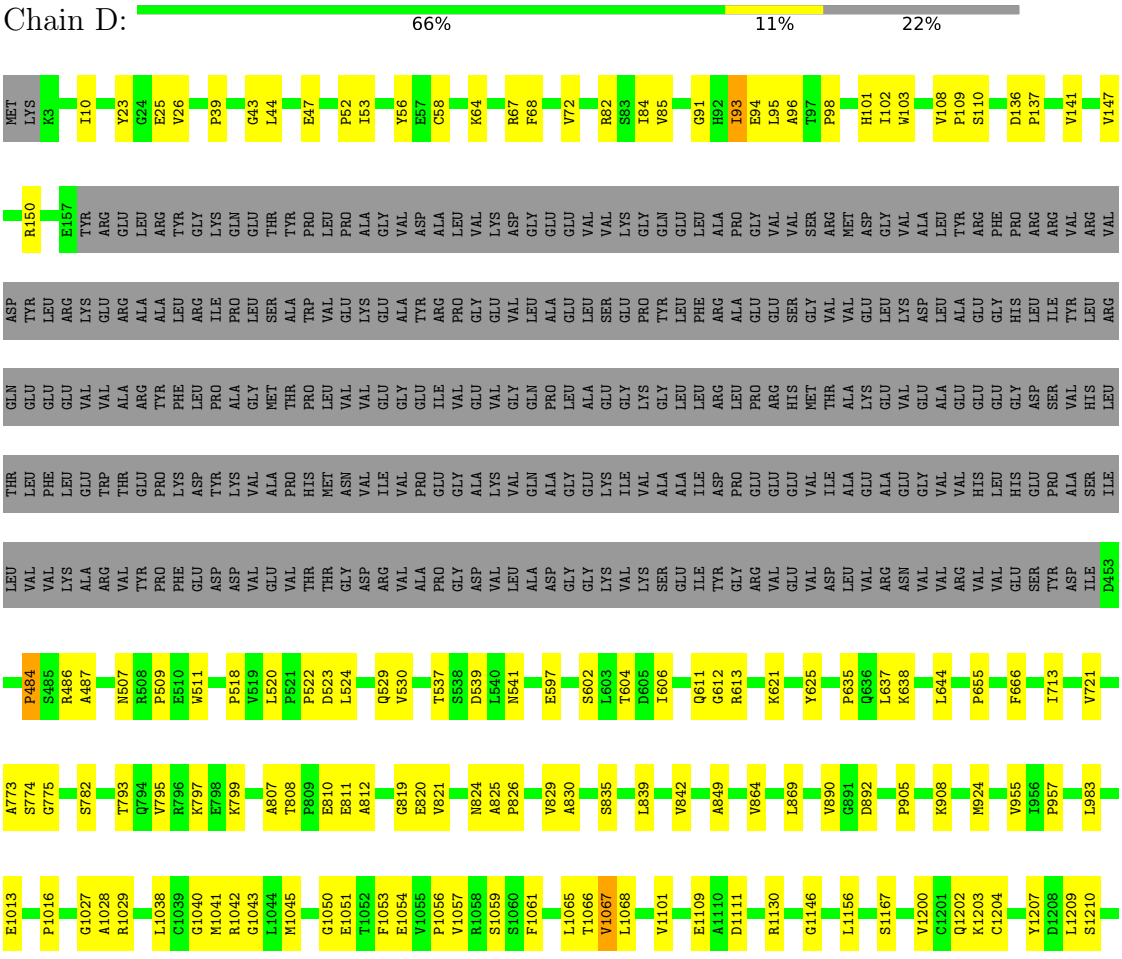
• Molecule 4: RNA POLYMERASE, BETA SUBUNIT

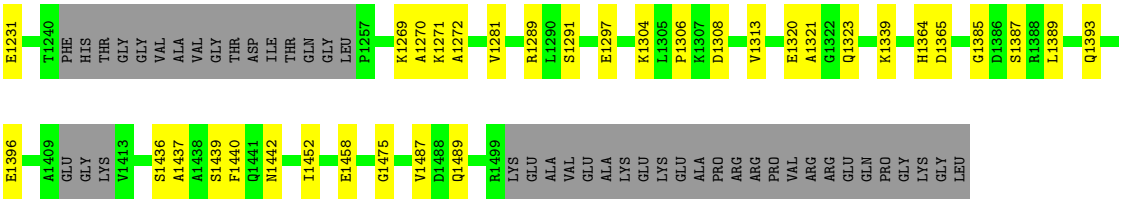
Chain C:



• Molecule 5: RNA POLYMERASE, BETA-PRIME SUBUNIT

Chain D:



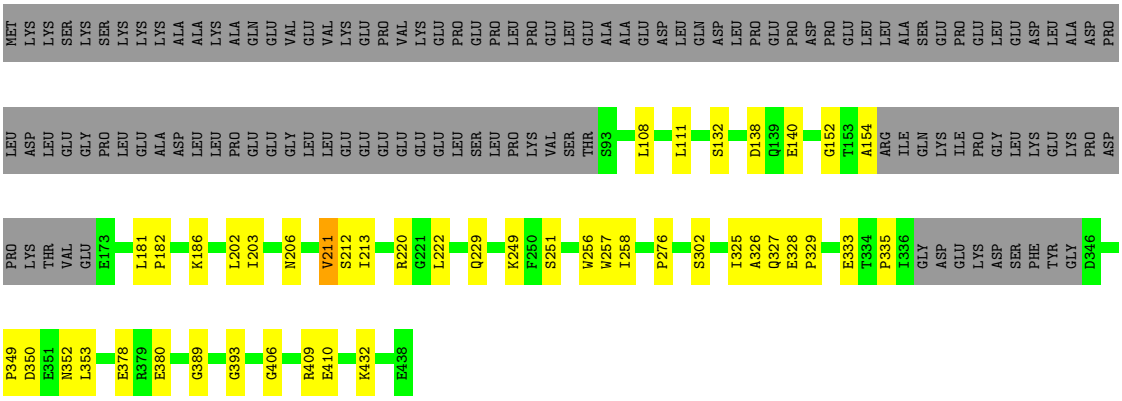


• Molecule 6: RNA POLYMERASE, OMEGA SUBUNIT

Chain E:

• Molecule 7: SIGMA FACTOR SIGA

Chain H:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 43 2 2	Depositor
Cell constants a, b, c, α, β, γ	180.95Å 180.95Å 523.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	80.00 – 6.50	Depositor
% Data completeness (in resolution range)	(Not available) (80.00-6.50)	Depositor
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
Refinement program		Depositor
R, R_{free}	(Not available) , (Not available)	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	10078	wwPDB-VP
Average B, all atoms (Å^2)	25.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	U	1.48	2/416 (0.5%)	3.85	36/622 (5.8%)
2	T	4.58	8/356 (2.2%)	3.82	23/532 (4.3%)
3	A	0.20	0/671	0.47	0/670
3	B	0.20	0/659	0.53	0/658
4	C	0.19	0/3246	0.41	0/3240
5	D	0.19	0/3545	0.41	0/3541
6	E	0.20	0/275	0.35	0/274
7	H	0.20	0/954	0.49	0/951
All	All	0.93	10/10122 (0.1%)	1.34	59/10488 (0.6%)

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
1	U	1	0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	T	174	DA	O3'-P	-48.81	1.02	1.61
2	T	185	DC	O3'-P	-48.24	1.03	1.61
2	T	172	DG	O3'-P	-39.32	1.14	1.61
2	T	175	DC	O3'-P	-23.63	1.32	1.61
1	U	26	DG	O3'-P	-23.04	1.33	1.61

The worst 5 of 59 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	U	28	DG	P-O3'-C3'	-45.80	64.74	119.70
1	U	30	DT	P-O3'-C3'	-41.74	69.61	119.70
2	T	171	DA	P-O3'-C3'	-36.19	76.27	119.70
2	T	173	DC	P-O3'-C3'	36.09	163.01	119.70
1	U	28	DG	OP2-P-O3'	30.75	172.85	105.20

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	U	28	DG	C3'

There are no planarity outliers.

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	U	382	0	212	23	0
2	T	327	0	184	28	0
3	A	672	0	241	2	0
3	B	660	0	238	2	0
4	C	3252	0	1199	10	0
5	D	3549	0	1291	23	0
6	E	276	0	96	0	0
7	H	957	0	339	6	0
8	D	1	0	0	0	0
9	D	2	0	0	0	0
All	All	10078	0	3800	93	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 93 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:T:174:DA:O3'	2:T:175:DC:P	1.02	1.41
2:T:185:DC:O3'	2:T:186:DT:P	1.03	1.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:U:29:DC:C1'	1:U:30:DT:H5'	1.49	1.40
1:U:28:DG:C1'	1:U:29:DC:H5'	1.24	1.38
2:T:174:DA:C3'	2:T:175:DC:P	2.34	1.15

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	A	222/314 (71%)	149 (67%)	46 (21%)	27 (12%)	0 5
3	B	218/314 (69%)	144 (66%)	47 (22%)	27 (12%)	0 5
4	C	1072/1118 (96%)	710 (66%)	251 (23%)	111 (10%)	0 8
5	D	1175/1524 (77%)	742 (63%)	294 (25%)	139 (12%)	0 6
6	E	90/99 (91%)	74 (82%)	11 (12%)	5 (6%)	2 18
7	H	313/438 (72%)	227 (72%)	50 (16%)	36 (12%)	0 6
All	All	3090/3807 (81%)	2046 (66%)	699 (23%)	345 (11%)	0 7

5 of 345 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	58	ILE
3	A	156	HIS
3	A	158	ILE
3	A	186	LEU
3	B	20	TYR

5.3.2 Protein sidechains (i)

There are no protein residues with a non-rotameric sidechain to report in this entry.

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

Of 3 ligands modelled in this entry, 3 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	T	5
1	U	1

The worst 5 of 6 chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	T	171:DA	O3'	172:DG	P	1.81

Continued on next page...

Continued from previous page...

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	U	26:DG	O3'	27:DT	P	1.33
1	T	175:DC	O3'	176:DA	P	1.32
1	T	172:DG	O3'	173:DC	P	1.14
1	T	185:DC	O3'	186:DT	P	1.03

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.