



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 07:18 PM UTC

PDB ID : 1H38 / pdb_00001h38
Title : Structure of a T7 RNA polymerase elongation complex at 2.9Å resolution
Authors : Tahirov, T.H.; Temyakov, D.; Anikin, M.; Patlan, V.; McAllister, W.T.; Vassilyev, D.G.; Yokoyama, S.
Deposited on : 2002-08-24
Resolution : 2.90 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity	:	4-5-2 with Phenix2.0
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

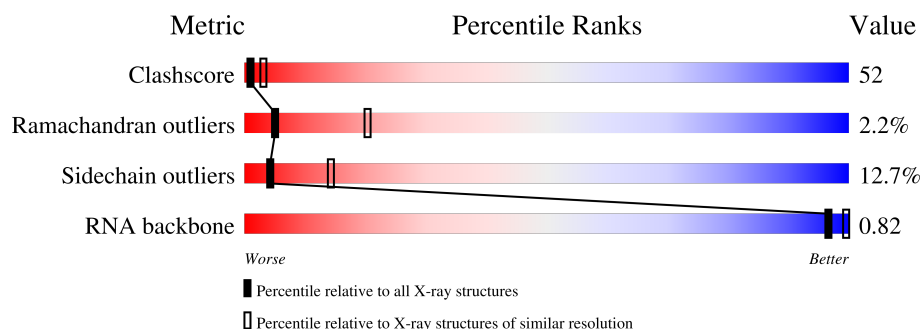
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.90 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	190562	2690 (2.90-2.90)
Ramachandran outliers	187476	2623 (2.90-2.90)
Sidechain outliers	187428	2625 (2.90-2.90)
RNA backbone	3983	1120 (3.10-2.70)

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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	883	36% 50% 10% ..
1	B	883	36% 49% 11% ..
1	C	883	37% 50% 9% ..
1	D	883	36% 51% 9% .
2	E	18	61% 17% 17% 6%
2	H	18	17% 39% 28% 17%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	K	18	
2	N	18	
3	F	12	
3	I	12	
3	L	12	
3	O	12	
4	G	10	
4	J	10	
4	M	10	
4	P	10	

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 30948 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	857	Total	C	N	O	S	0	0	0
			6746	4296	1173	1242	35			
1	B	857	Total	C	N	O	S	0	0	0
			6746	4296	1173	1242	35			
1	C	857	Total	C	N	O	S	0	0	0
			6746	4296	1173	1242	35			
1	D	857	Total	C	N	O	S	0	0	0
			6746	4296	1173	1242	35			

- Molecule 2 is a DNA chain called 5'-D(*GP*GP*GP*AP*AP*TP*CP*GP*AP*CP *AP*TP*CP*GP*CP*CP*GP*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	17	Total	C	N	O	P	0	0	0
			345	164	67	98	16			
2	H	18	Total	C	N	O	P	0	0	0
			367	174	72	104	17			
2	K	17	Total	C	N	O	P	0	0	0
			345	164	67	98	16			
2	N	17	Total	C	N	O	P	0	0	0
			345	164	67	98	16			

- Molecule 3 is a RNA chain called 5'-R(*AP*AP*CP*UP*GP*CP*GP*GP*CP*GP *AP*U)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	8	Total	C	N	O	P	0	0	0
			171	77	33	54	7			
3	I	8	Total	C	N	O	P	0	0	0
			171	77	33	54	7			
3	L	8	Total	C	N	O	P	0	0	0
			171	77	33	54	7			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	O	8	Total	C	N	O	P	0	0	0
			171	77	33	54	7			

- Molecule 4 is a DNA chain called 5'-D(*GP*TP*CP*GP*AP*TP*TP*CP*CP*CP)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	G	9	Total	C	N	O	P	0	0	0
			179	87	30	54	8			
4	J	9	Total	C	N	O	P	0	0	0
			179	87	30	54	8			
4	M	9	Total	C	N	O	P	0	0	0
			179	87	30	54	8			
4	P	9	Total	C	N	O	P	0	0	0
			179	87	30	54	8			

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	310	Total	O	0	0
			310	310		
5	B	352	Total	O	0	0
			352	352		
5	C	185	Total	O	0	0
			185	185		
5	D	177	Total	O	0	0
			177	177		
5	E	13	Total	O	0	0
			13	13		
5	F	16	Total	O	0	0
			16	16		
5	G	12	Total	O	0	0
			12	12		
5	H	20	Total	O	0	0
			20	20		
5	I	11	Total	O	0	0
			11	11		
5	J	9	Total	O	0	0
			9	9		
5	K	17	Total	O	0	0
			17	17		
5	L	5	Total	O	0	0
			5	5		

Continued on next page...

Continued from previous page...

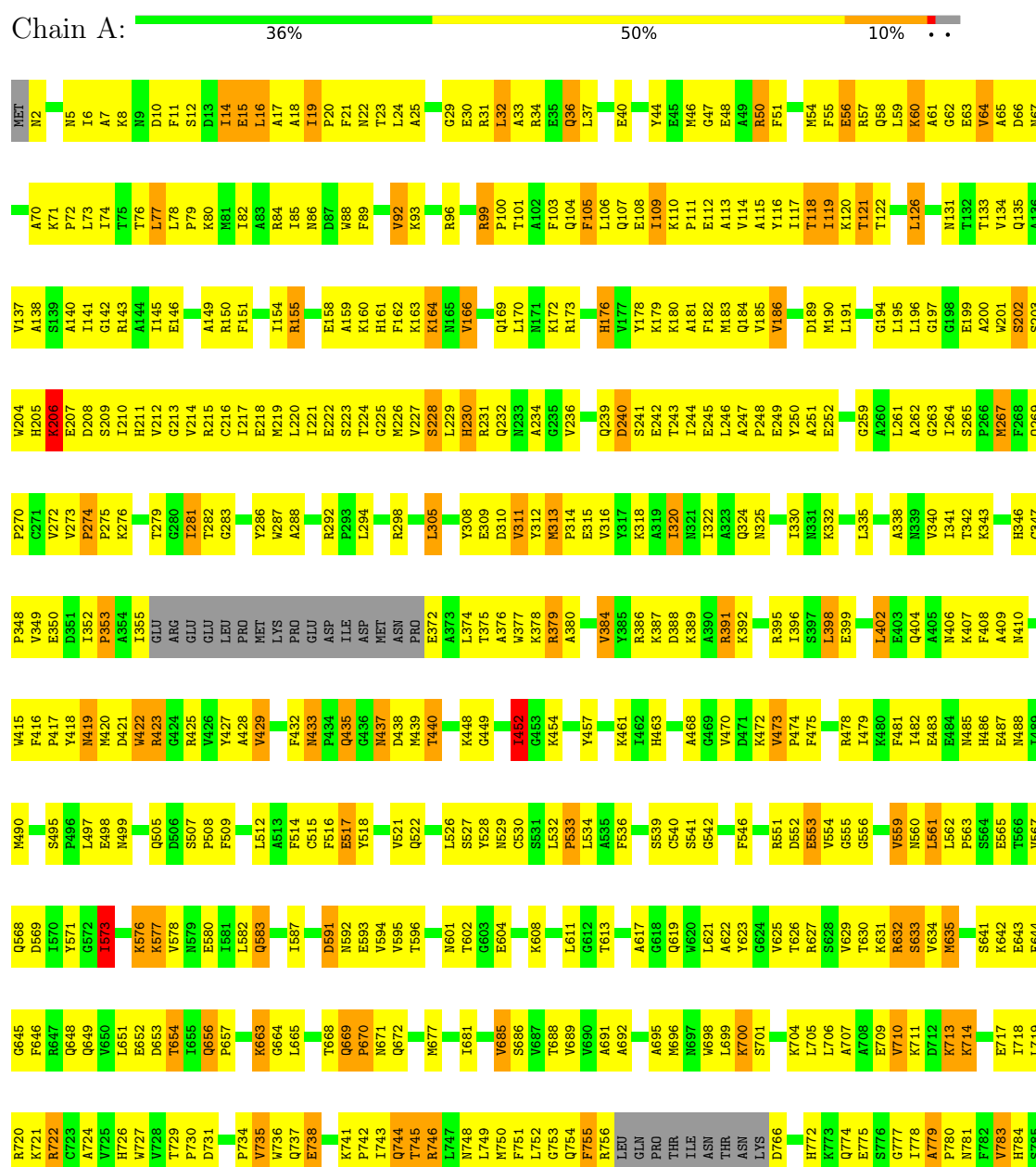
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	M	8	Total 8	O 8	0	0
5	N	12	Total 12	O 12	0	0
5	O	6	Total 6	O 6	0	0
5	P	9	Total 9	O 9	0	0

3 Residue-property plots

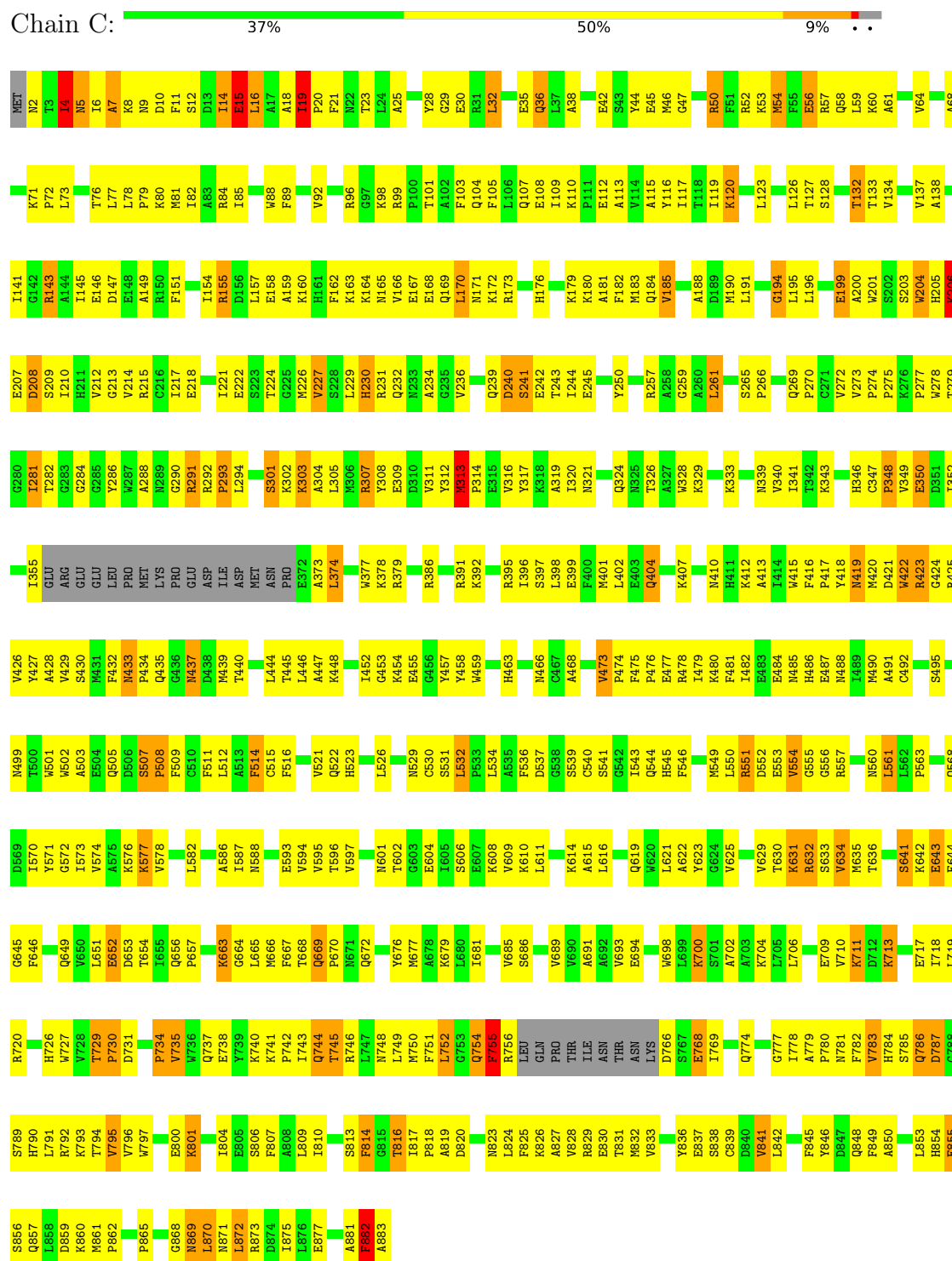
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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.

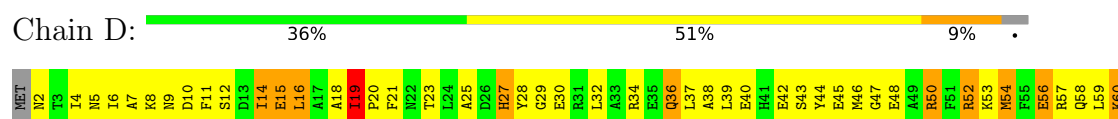
• Molecule 1: DNA-DIRECTED RNA POLYMERASE

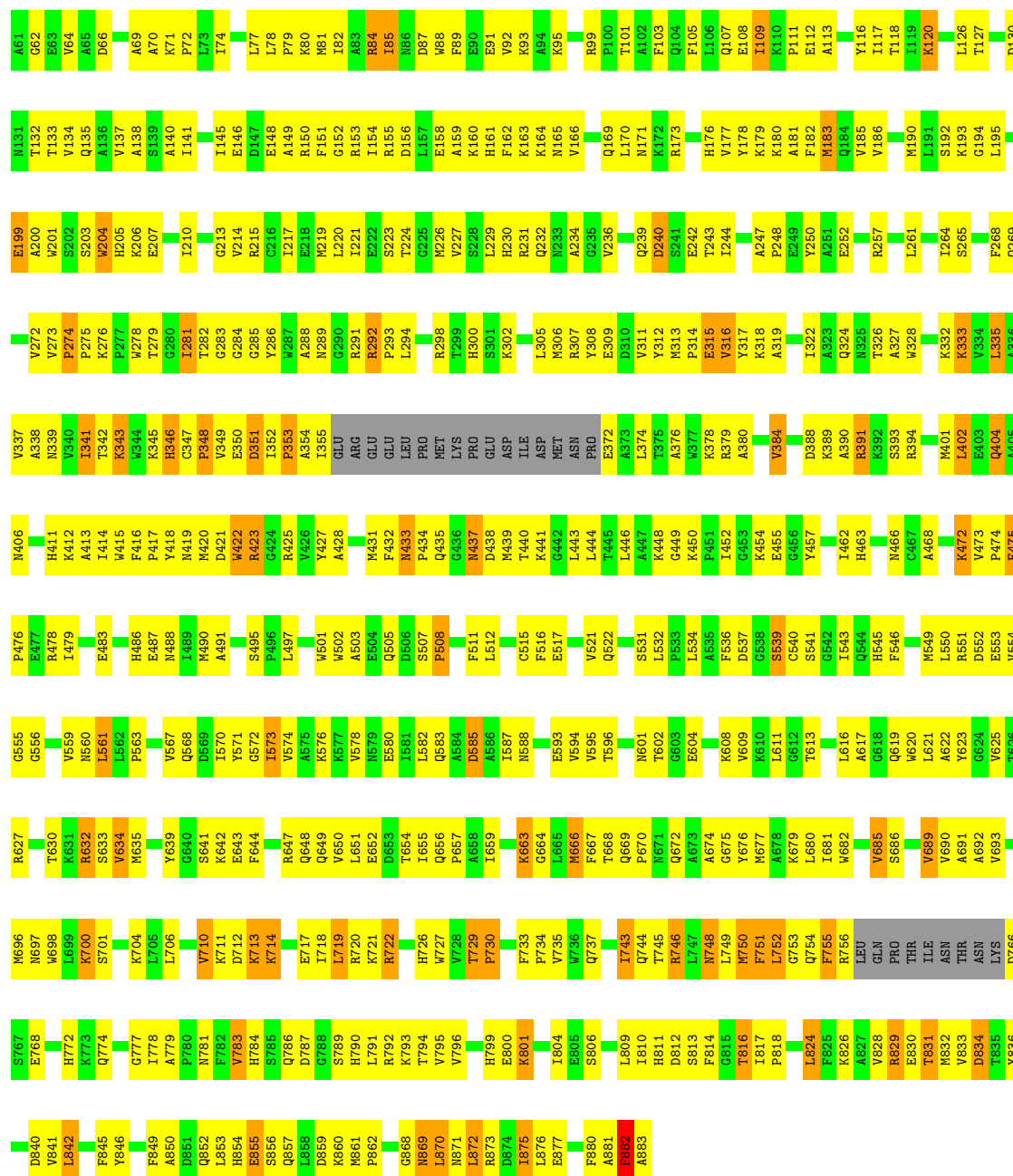


● Molecule 1: DNA-DIRECTED RNA POLYMERASE



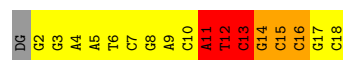
● Molecule 1: DNA-DIRECTED RNA POLYMERASE





• Molecule 2: 5'-D(*GP*GP*GP*AP*AP*TP*CP*GP*AP*CP *AP*TP*CP*GP*CP*CP*GP*C)-3'

Chain E:



• Molecule 2: 5'-D(*GP*GP*GP*AP*AP*TP*CP*GP*AP*CP *AP*TP*CP*GP*CP*CP*GP*C)-3'

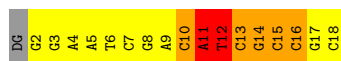
Chain H:





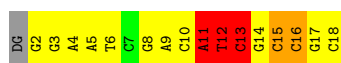
• Molecule 2: 5'-D(*GP*GP*GP*AP*AP*TP*CP*GP*AP*CP *AP*TP*CP*GP*CP*CP*GP*C)-3'

Chain K: 56% 28% 11% 6%



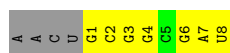
• Molecule 2: 5'-D(*GP*GP*GP*AP*AP*TP*CP*GP*AP*CP *AP*TP*CP*GP*CP*CP*GP*C)-3'

Chain N: 6% 61% 11% 17% 6%



• Molecule 3: 5'-R(*AP*AP*CP*UP*GP*CP*GP*GP*CP*GP *AP*U)-3'

Chain F: 8% 58% 33%



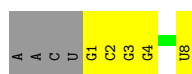
• Molecule 3: 5'-R(*AP*AP*CP*UP*GP*CP*GP*GP*CP*GP *AP*U)-3'

Chain I: 17% 50% 33%



• Molecule 3: 5'-R(*AP*AP*CP*UP*GP*CP*GP*GP*CP*GP *AP*U)-3'

Chain L: 25% 42% 33%



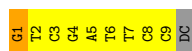
• Molecule 3: 5'-R(*AP*AP*CP*UP*GP*CP*GP*GP*CP*GP *AP*U)-3'

Chain O: 25% 42% 33%



• Molecule 4: 5'-D(*GP*TP*CP*GP*AP*TP*TP*CP*CP*CP)-3'

Chain G: 80% 10% 10%



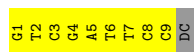
- Molecule 4: 5'-D(*GP*TP*CP*GP*AP*TP*TP*CP*CP*CP)-3'

Chain J:  20% 70% 10%



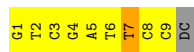
- Molecule 4: 5'-D(*GP*TP*CP*GP*AP*TP*TP*CP*CP*CP)-3'

Chain M:  90% 10%



- Molecule 4: 5'-D(*GP*TP*CP*GP*AP*TP*TP*CP*CP*CP)-3'

Chain P:  80% 10% 10%



4 Data and refinement statistics

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

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	79.91Å 84.97Å 202.00Å 90.36° 92.97° 109.94°	Depositor
Resolution (Å)	39.93 – 2.90	Depositor
% Data completeness (in resolution range)	98.0 (39.93-2.90)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.236 , 0.284	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	30948	wwPDB-VP
Average B, all atoms (Å ²)	71.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

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	1.10	21/6897 (0.3%)	1.12	22/9329 (0.2%)
1	B	1.12	18/6897 (0.3%)	1.15	22/9329 (0.2%)
1	C	0.76	3/6897 (0.0%)	1.00	8/9329 (0.1%)
1	D	0.69	1/6897 (0.0%)	0.97	13/9329 (0.1%)
2	E	0.53	0/387	1.13	3/595 (0.5%)
2	H	0.52	0/412	1.10	3/634 (0.5%)
2	K	0.50	0/387	1.09	2/595 (0.3%)
2	N	0.49	0/387	1.08	3/595 (0.5%)
3	F	0.67	0/191	0.75	0/297
3	I	0.68	0/191	0.78	0/297
3	L	0.53	0/191	0.75	0/297
3	O	0.45	0/191	0.69	0/297
4	G	0.50	0/199	0.82	0/305
4	J	0.47	0/199	0.95	0/305
4	M	0.42	0/199	0.88	0/305
4	P	0.46	0/199	1.00	0/305
All	All	0.91	43/30721 (0.1%)	1.05	76/42143 (0.2%)

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	E	0	6
2	H	0	8
2	K	0	7
2	N	0	6
4	G	0	1
4	J	0	1
4	P	0	1
All	All	0	30

The worst 5 of 43 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	573	ILE	CA-CB	12.05	1.71	1.54
1	A	783	VAL	CA-CB	-9.90	1.43	1.54
1	B	778	ILE	CA-CB	9.19	1.65	1.54
1	A	630	THR	CA-CB	-8.85	1.42	1.54
1	A	654	THR	CA-CB	8.44	1.67	1.53

The worst 5 of 76 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	828	VAL	CB-CA-C	-8.20	101.07	112.14
2	H	11	DA	N9-C1'-C2'	-8.05	101.42	113.50
1	B	495	SER	CA-C-N	7.13	126.97	119.05
1	B	495	SER	C-N-CA	7.13	126.97	119.05
2	E	11	DA	N9-C1'-C2'	-7.09	102.86	113.50

There are no chirality outliers.

5 of 30 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	E	11	DA	Sidechain
2	E	12	DT	Sidechain
2	E	13	DC	Sidechain
2	E	14	DG	Sidechain
2	E	15	DC	Sidechain

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	6746	0	6708	752	0
1	B	6746	0	6708	750	0
1	C	6746	0	6708	707	0
1	D	6746	0	6708	686	0
2	E	345	0	191	29	0
2	H	367	0	202	28	0
2	K	345	0	191	38	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	N	345	0	191	21	0
3	F	171	0	89	7	0
3	I	171	0	89	11	0
3	L	171	0	89	9	0
3	O	171	0	89	5	0
4	G	179	0	104	16	0
4	J	179	0	104	8	0
4	M	179	0	104	19	0
4	P	179	0	104	11	0
5	A	310	0	0	104	0
5	B	352	0	0	131	0
5	C	185	0	0	78	0
5	D	177	0	0	108	0
5	E	13	0	0	2	0
5	F	16	0	0	2	0
5	G	12	0	0	6	0
5	H	20	0	0	6	0
5	I	11	0	0	4	0
5	J	9	0	0	4	0
5	K	17	0	0	7	0
5	L	5	0	0	2	0
5	M	8	0	0	1	0
5	N	12	0	0	4	0
5	O	6	0	0	0	0
5	P	9	0	0	3	0
All	All	30948	0	28379	3042	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 52.

The worst 5 of 3042 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)
1:A:668:THR:HG22	1:A:669:GLN:NE2	1.51	1.24
1:C:428:ALA:H	1:C:435:GLN:NE2	1.44	1.15
1:B:50:ARG:HG2	1:B:50:ARG:HH11	1.01	1.14
1:C:133:THR:HA	1:C:243:THR:HG22	1.28	1.10
1:A:120:LYS:HE3	1:A:752:LEU:HD21	1.31	1.10

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	851/883 (96%)	740 (87%)	94 (11%)	17 (2%)	6	22
1	B	851/883 (96%)	736 (86%)	98 (12%)	17 (2%)	6	22
1	C	851/883 (96%)	737 (87%)	93 (11%)	21 (2%)	4	17
1	D	851/883 (96%)	735 (86%)	96 (11%)	20 (2%)	4	18
All	All	3404/3532 (96%)	2948 (87%)	381 (11%)	75 (2%)	5	20

5 of 75 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	194	GLY
1	A	199	GLU
1	A	539	SER
1	A	663	LYS
1	A	755	PHE

5.3.2 Protein sidechains ⓘ

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	703/729 (96%)	614 (87%)	89 (13%)	4	14
1	B	703/729 (96%)	603 (86%)	100 (14%)	3	11
1	C	703/729 (96%)	617 (88%)	86 (12%)	5	16
1	D	703/729 (96%)	620 (88%)	83 (12%)	5	17
All	All	2812/2916 (96%)	2454 (87%)	358 (13%)	4	14

5 of 358 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	473	VAL
1	D	84	ARG
1	C	577	LYS
1	C	754	GLN
1	D	335	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 150 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	171	ASN
1	D	786	GLN
1	D	239	GLN
1	D	463	HIS
1	B	239	GLN

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	F	7/12 (58%)	0	0
3	I	7/12 (58%)	0	0
3	L	7/12 (58%)	0	0
3	O	7/12 (58%)	0	0
All	All	28/48 (58%)	0	0

There are no RNA backbone outliers to report.

There are no RNA pucker outliers to report.

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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](#)

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.