



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 05:34 am GMT

PDB ID : 4OIR
Title : Crystal structure of Thermus thermophilus RNA polymerase transcription initiation complex soaked with GE23077 and rifamycin SV
Authors : Zhang, Y.; Ebright, R.H.; Arnold, E.
Deposited on : 2014-01-20
Resolution : 3.10 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

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc28949

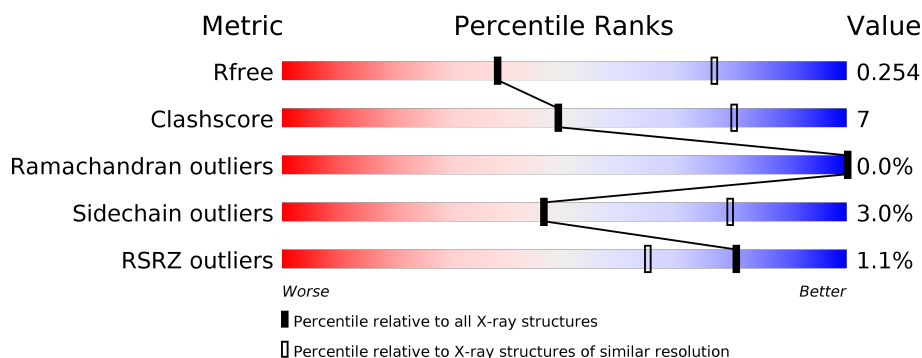
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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}	100719	1001 (3.12-3.08)
Clashscore	112137	1099 (3.12-3.08)
Ramachandran outliers	110173	1057 (3.12-3.08)
Sidechain outliers	110143	1057 (3.12-3.08)
RSRZ outliers	101464	1006 (3.12-3.08)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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 $\leq 5\%$. 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	305	<div> <div style="width: 59%;"></div> <div style="width: 16%;"></div> <div style="width: 24%;"></div> </div>
1	B	305	<div> <div style="width: 56%;"></div> <div style="width: 18%;"></div> <div style="width: 26%;"></div> </div>
2	C	1119	<div> <div style="width: 81%;"></div> <div style="width: 17%;"></div> <div style="width: 2%;"></div> </div>
3	D	1524	<div> <div style="width: 79%;"></div> <div style="width: 18%;"></div> <div style="width: 3%;"></div> </div>
4	E	99	<div> <div style="width: 82%;"></div> <div style="width: 12%;"></div> <div style="width: 5%;"></div> </div>
5	F	443	<div> <div style="width: 64%;"></div> <div style="width: 13%;"></div> <div style="width: 22%;"></div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
6	G	21	<div><div></div><div></div><div></div><div>43%</div><div>33%</div><div>24%</div></div>
7	H	27	<div><div></div><div></div><div></div><div>22%</div><div>59%</div><div>7%</div><div>11%</div></div>
8	I	7	<div><div></div><div></div><div></div><div>57%</div><div>43%</div></div>

2 Entry composition

There are 13 unique types of molecules in this entry. The entry contains 28730 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	231	Total	C	N	O	S	0	0	0
			1809	1155	315	337	2			
1	B	227	Total	C	N	O	S	0	0	0
			1789	1143	310	334	2			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1112	Total	C	N	O	S	0	0	0
			8774	5550	1565	1635	24			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1485	Total	C	N	O	S	0	1	0
			11739	7441	2069	2193	36			

- 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
			758	483	132	139	4			

- Molecule 5 is a protein called DNA directed RNA polymerase sigma factor A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	346	Total	C	N	O	S	0	0	0
			2807	1770	509	524	4			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	-19	MET	-	EXPRESSION TAG	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

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	16	Total	C	N	O	P	0	0	0
			328	156	63	94	15			

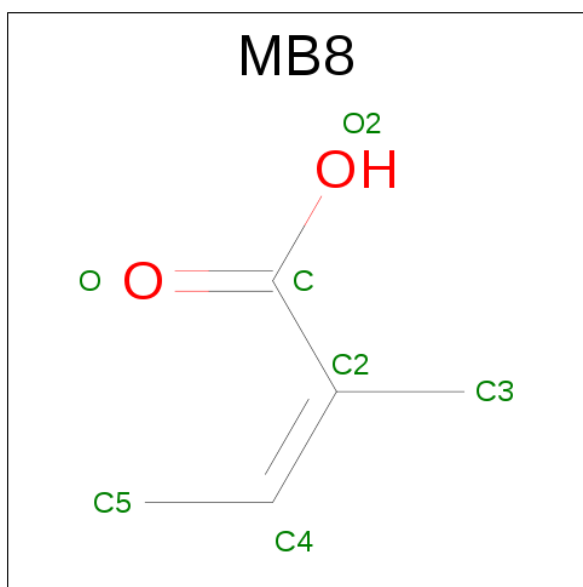
- Molecule 7 is a DNA chain called 5'-D(*TP*AP*TP*AP*AP*TP*GP*GP*GP*AP*GP*C P*TP*GP*TP*CP*AP*CP*GP*GP*AP*TP*GP*CP*AP*GP*G)-3'.

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

- Molecule 8 is a protein called GE23077.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
8	I	7	Total	C	N	O	0	0	0
			50	26	9	15			

- Molecule 9 is (2Z)-2-METHYLBUT-2-ENOIC ACID (three-letter code: MB8) (formula: C₅H₈O₂).

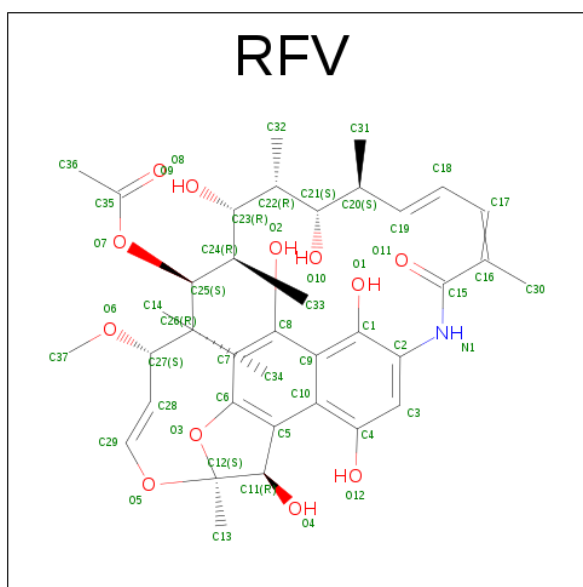


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	I	1	Total	C	O	0	0
			2	1	1		

- Molecule 10 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

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

- Molecule 11 is RIFAMYCIN SV (three-letter code: RFV) (formula: C₃₇H₄₉NO₁₂).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
11	C	1	Total	C	N	O	0	0
			50	37	1	12		

- Molecule 12 is ZINC ION (three-letter code: ZN) (formula: Zn).

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

- Molecule 13 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	A	8	Total	O	0	0
			8	8		
13	B	6	Total	O	0	0
			6	6		
13	C	46	Total	O	0	0
			46	46		
13	D	46	Total	O	0	0
			46	46		
13	E	3	Total	O	0	0
			3	3		
13	F	9	Total	O	0	0
			9	9		
13	G	4	Total	O	0	0
			4	4		

Continued on next page...

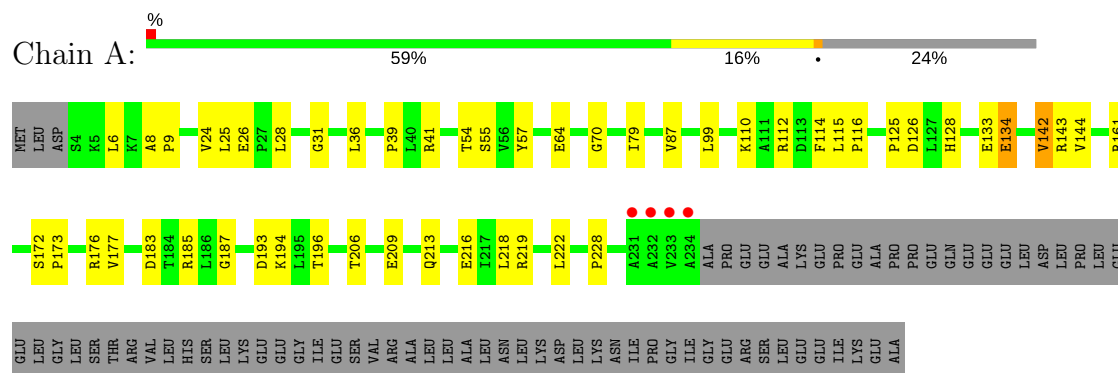
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
13	H	1	Total	O	0	0
			1	1		

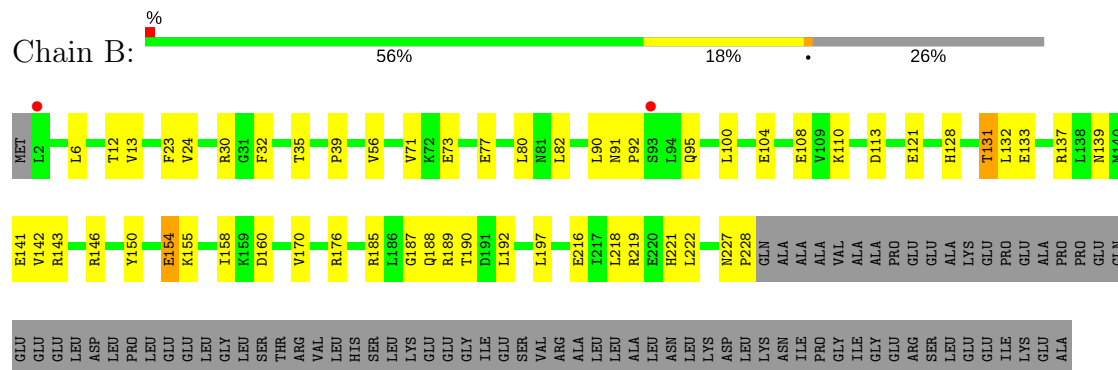
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

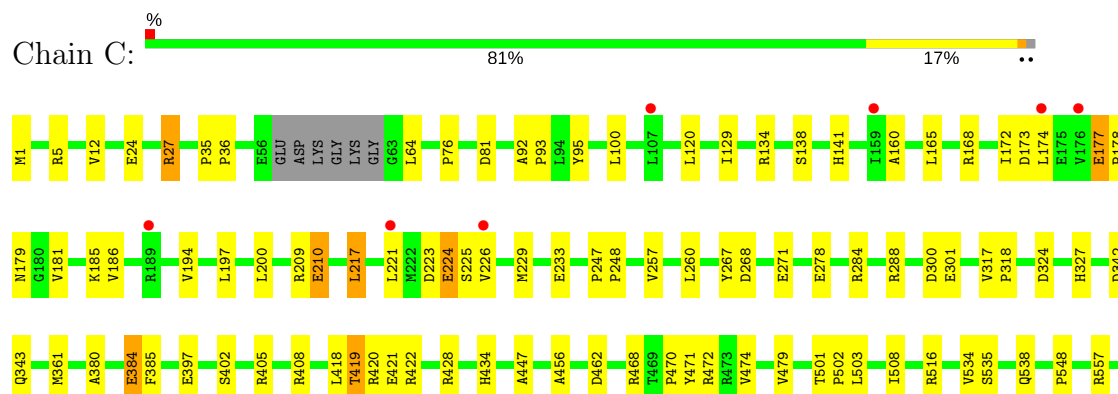
- Molecule 1: DNA-directed RNA polymerase subunit alpha

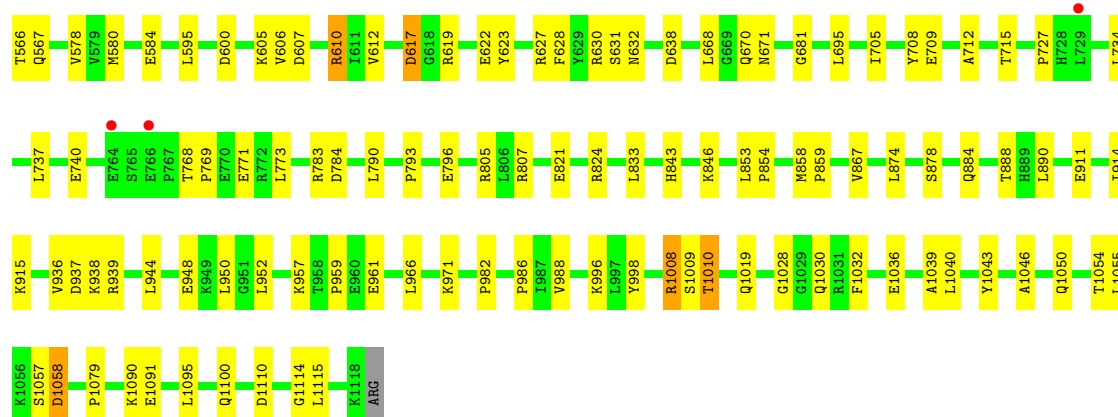


- Molecule 1: DNA-directed RNA polymerase subunit alpha

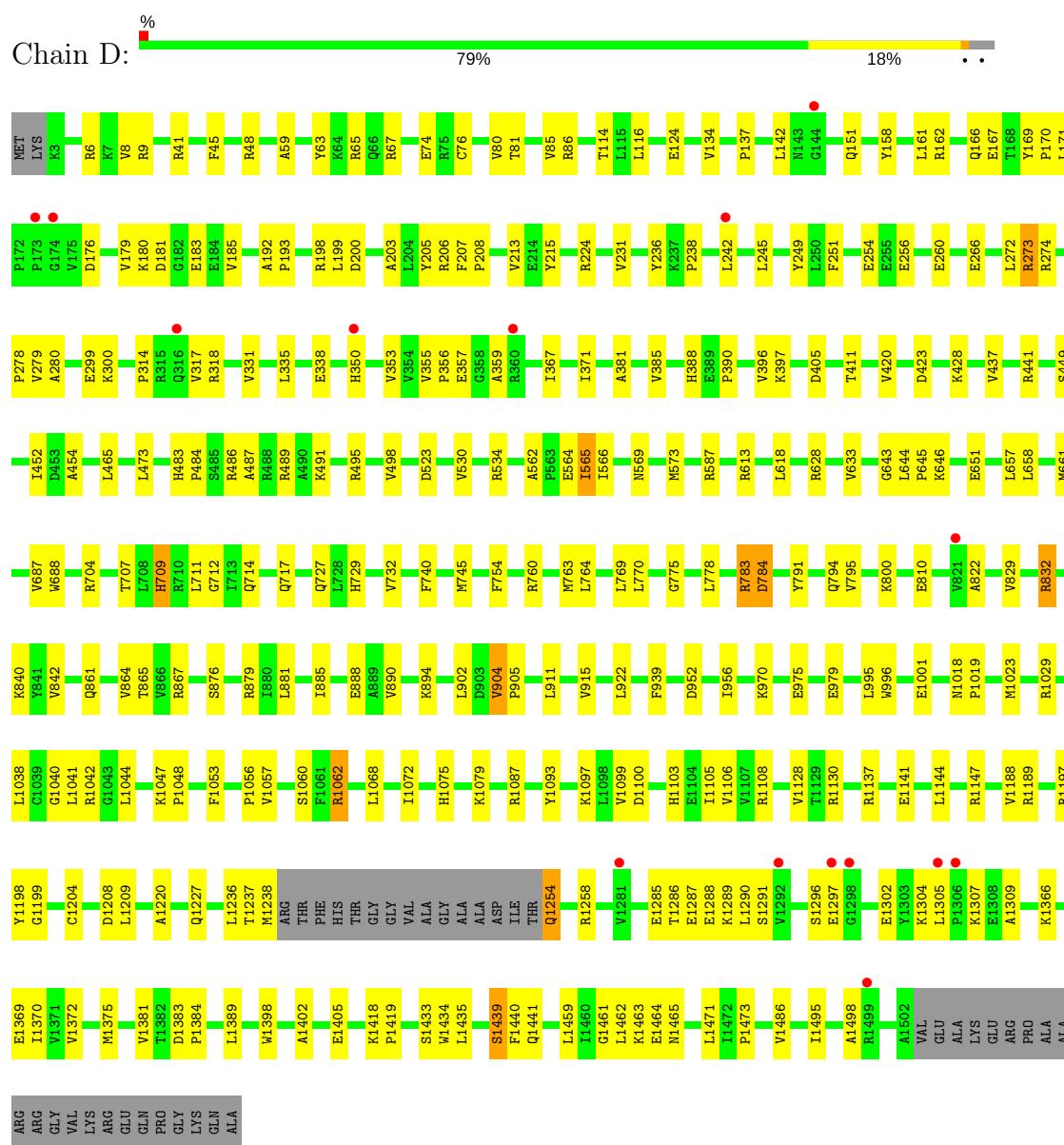


- Molecule 2: DNA-directed RNA polymerase subunit beta






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



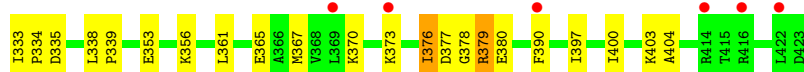
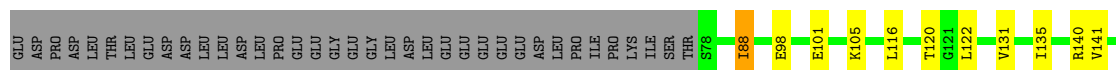
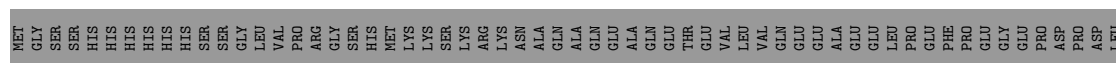
● Molecule 4: DNA-directed RNA polymerase subunit omega

Chain E:  82% 12% 5%



- Molecule 5: DNA directed RNA polymerase sigma factor A

Chain F:  2% 64% 13% 22%



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

Chain G:  43% 33% 24%



- Molecule 7: 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:  22% 59% 7% 11%



- Molecule 8: GE23077

Chain I:  57% 43%



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	183.70Å 105.01Å 295.36Å 90.00° 98.79° 90.00°	Depositor
Resolution (Å)	47.50 – 3.10 47.50 – 3.11	Depositor EDS
% Data completeness (in resolution range)	82.6 (47.50-3.10) 82.7 (47.50-3.11)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	0.14	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.71 (at 3.12Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, R_{free}	0.206 , 0.253 0.206 , 0.254	Depositor DCC
R_{free} test set	4156 reflections (5.00%)	DCC
Wilson B-factor (Å ²)	63.5	Xtriage
Anisotropy	0.074	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 25.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.018 for $1/2^*h-3/2^*k,-1/2^*h-1/2^*k,-1/2^*h+1/2^*k-l$ 0.015 for $1/2^*h+3/2^*k,1/2^*h-1/2^*k,-1/2^*h-1/2^*k-l$	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	28730	wwPDB-VP
Average B, all atoms (Å ²)	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.18% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: FGL, ZN, RFV, 2TL, DVA, MG, 2RA, DSN, MB8, 0QZ, R2T

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	0.24	0/1841	0.46	0/2504
1	B	0.23	0/1821	0.46	0/2476
2	C	0.24	0/8941	0.45	0/12092
3	D	0.24	0/11948	0.45	0/16152
4	E	0.24	0/772	0.41	0/1040
5	F	0.23	0/2852	0.40	0/3837
6	G	0.48	0/368	1.06	2/567 (0.4%)
7	H	0.50	0/556	1.14	3/858 (0.3%)
All	All	0.25	0/29099	0.49	5/39526 (0.0%)

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
8	I	0	1

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	G	5	DC	O4'-C4'-C3'	-9.49	100.30	106.00
7	H	17	DA	O4'-C1'-N9	7.70	113.39	108.00
7	H	23	DG	C4'-C3'-C2'	-5.43	98.22	103.10
7	H	16	DC	O4'-C1'-N1	5.04	111.53	108.00
6	G	5	DC	C4'-C3'-C2'	-5.01	98.59	103.10

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
8	I	5	2TL	Peptide

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	1809	0	1863	32	0
1	B	1789	0	1841	33	0
2	C	8774	0	8877	129	0
3	D	11739	0	11975	173	0
4	E	758	0	770	10	0
5	F	2807	0	2882	46	0
6	G	328	0	181	5	0
7	H	495	0	272	12	0
8	I	50	0	37	3	0
9	I	2	0	0	0	0
10	B	1	0	0	0	0
10	D	2	0	0	0	0
10	F	1	0	0	0	0
11	C	50	0	48	2	0
12	D	2	0	0	0	0
13	A	8	0	0	0	0
13	B	6	0	0	0	0
13	C	46	0	0	1	0
13	D	46	0	0	4	0
13	E	3	0	0	0	0
13	F	9	0	0	2	0
13	G	4	0	0	0	0
13	H	1	0	0	0	0
All	All	28730	0	28746	392	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 392 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:F:321:ILE:HD11	5:F:329:TYR:HA	1.69	0.74
2:C:12:VAL:HG21	2:C:472:ARG:HD3	1.69	0.74
3:D:256:GLU:HG3	3:D:300:LYS:HG3	1.69	0.74
11:C:1201:RFV:H51	11:C:1201:RFV:H48	1.33	0.72
2:C:428:ARG:NH2	2:C:447:ALA:O	2.22	0.72

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	229/305 (75%)	226 (99%)	3 (1%)	0	100	100
1	B	225/305 (74%)	223 (99%)	2 (1%)	0	100	100
2	C	1108/1119 (99%)	1080 (98%)	28 (2%)	0	100	100
3	D	1482/1524 (97%)	1450 (98%)	31 (2%)	1 (0%)	55	88
4	E	92/99 (93%)	90 (98%)	2 (2%)	0	100	100
5	F	344/443 (78%)	339 (98%)	5 (2%)	0	100	100
All	All	3480/3795 (92%)	3408 (98%)	71 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	565	ILE

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	200/264 (76%)	196 (98%)	4 (2%)	60	86
1	B	200/264 (76%)	190 (95%)	10 (5%)	28	65
2	C	936/941 (100%)	906 (97%)	30 (3%)	44	78
3	D	1253/1279 (98%)	1219 (97%)	34 (3%)	50	81
4	E	82/88 (93%)	81 (99%)	1 (1%)	75	91
5	F	301/388 (78%)	290 (96%)	11 (4%)	39	75
All	All	2972/3224 (92%)	2882 (97%)	90 (3%)	46	79

5 of 90 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	1010	THR
3	D	273	ARG
5	F	295	MET
2	C	1057	SER
3	D	80	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	696	HIS
3	D	1195	GLN
3	D	724	GLN
3	D	350	HIS
3	D	714	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

7 non-standard protein/DNA/RNA residues are modelled in this entry.

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$
8	2RA	I	1	9,8	4,5,6	1.28	1 (25%)	1,5,7	1.82	0
8	DSN	I	2	8	5,5,6	1.62	1 (20%)	1,5,7	1.44	0
8	DVA	I	3	8	6,6,7	1.37	1 (16%)	6,7,9	1.34	1 (16%)
8	R2T	I	4	8	10,10,11	2.44	3 (30%)	8,13,15	1.22	1 (12%)
8	2TL	I	5	8	6,6,7	0.94	0	6,7,9	1.10	1 (16%)
8	0QZ	I	6	8	5,5,6	1.43	1 (20%)	4,5,7	1.39	1 (25%)
8	FGL	I	7	8	1,6,7	0.75	0	2,7,9	2.20	1 (50%)

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
8	2RA	I	1	9,8	-	0/2/4/6	0/0/0/0
8	DSN	I	2	8	-	0/2/4/6	0/0/0/0
8	DVA	I	3	8	-	0/5/6/8	0/0/0/0
8	R2T	I	4	8	-	0/11/14/16	0/0/0/0
8	2TL	I	5	8	-	0/4/6/8	0/0/0/0
8	0QZ	I	6	8	-	0/2/4/6	0/0/0/0
8	FGL	I	7	8	-	0/0/6/8	0/0/0/0

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	I	6	0QZ	OB-CA	-2.99	1.38	1.43
8	I	4	R2T	OB1-CB	-2.30	1.37	1.43
8	I	1	2RA	CA-C	2.36	1.53	1.50
8	I	3	DVA	CA-C	3.16	1.54	1.50
8	I	2	DSN	CA-C	3.47	1.54	1.50

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	I	7	FGL	O-C-CA	-3.07	119.91	124.40
8	I	3	DVA	O-C-CA	-2.98	118.20	125.15
8	I	6	0QZ	O-C-CA	-2.48	119.14	125.06

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	I	5	2TL	O-C-CA	-2.13	120.18	125.15
8	I	4	R2T	C-CA-N	2.37	114.63	109.86

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	I	4	R2T	2	0
8	I	7	FGL	1	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 6 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	RFV	C	1201	-	51,53,53	3.46	10 (19%)	72,80,80	2.35	14 (19%)
9	MB8	I	101	8	0,1,6	0.00	-	0,0,7	0.00	-

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	RFV	C	1201	-	-	0/55/70/70	0/1/4/4

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
9	MB8	I	101	8	-	0/0/0/6	0/0/0/0

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
11	C	1201	RFV	C12-C11	-20.64	1.40	1.55
11	C	1201	RFV	O7-C25	-4.95	1.37	1.44
11	C	1201	RFV	O9-C23	-2.24	1.37	1.43
11	C	1201	RFV	C5-C10	2.20	1.46	1.42
11	C	1201	RFV	O3-C12	2.77	1.49	1.45

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
11	C	1201	RFV	C30-C16-C17	-5.92	107.83	123.44
11	C	1201	RFV	C13-C12-C11	-3.26	110.02	117.83
11	C	1201	RFV	O3-C6-C5	-3.16	109.81	113.28
11	C	1201	RFV	C33-C24-C25	-2.97	105.95	111.43
11	C	1201	RFV	C20-C19-C18	-2.24	120.81	126.06

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
11	C	1201	RFV	2	0

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, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å ²)	Q<0.9
1	A	231/305 (75%)	-0.43	4 (1%) 70 49	31, 50, 79, 128	0
1	B	227/305 (74%)	-0.36	2 (0%) 84 69	35, 65, 92, 125	0
2	C	1112/1119 (99%)	-0.35	10 (0%) 84 69	15, 44, 109, 139	0
3	D	1485/1524 (97%)	-0.28	15 (1%) 82 67	15, 50, 111, 156	0
4	E	94/99 (94%)	-0.40	0 100 100	27, 55, 99, 107	0
5	F	346/443 (78%)	-0.17	7 (2%) 65 44	24, 66, 122, 138	0
6	G	16/21 (76%)	-0.32	0 100 100	66, 101, 179, 183	0
7	H	24/27 (88%)	-0.24	0 100 100	58, 110, 171, 185	0
8	I	0/7	-	-	-	-
All	All	3535/3850 (91%)	-0.31	38 (1%) 80 65	15, 52, 112, 185	0

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	232	ALA	4.0
2	C	766	GLU	3.2
3	D	1281	VAL	2.9
1	A	233	VAL	2.9
5	F	149	GLU	2.9

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

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
8	FGL	I	7	7/8	0.93	0.15	-	19,25,27,29	0
8	2TL	I	5	7/8	0.98	0.18	-	21,21,24,29	0
8	2RA	I	1	6/7	0.96	0.11	-	24,25,29,34	0
8	DVA	I	3	7/8	0.97	0.17	-	13,17,21,22	0
8	DSN	I	2	6/7	0.95	0.17	-	20,21,29,31	0
8	0QZ	I	6	6/7	0.97	0.15	-	23,25,25,28	0
8	R2T	I	4	11/12	0.95	0.16	-	21,25,32,33	0

6.3 Carbohydrates [i](#)

There are no carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
11	RFV	C	1201	50/50	0.95	0.24	0.86	19,30,47,53	0
10	MG	B	401	1/1	0.96	0.25	0.32	25,25,25,25	0
10	MG	F	2001	1/1	0.96	0.17	-0.06	55,55,55,55	0
12	ZN	D	2002	1/1	0.97	0.08	-0.81	79,79,79,79	0
12	ZN	D	2001	1/1	0.97	0.17	-0.88	43,43,43,43	0
10	MG	D	2004	1/1	0.77	0.18	-	50,50,50,50	0
10	MG	D	2003	1/1	0.97	0.17	-	14,14,14,14	0
9	MB8	I	101	2/7	0.96	0.13	-	26,26,26,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.