Full wwPDB X-ray Structure Validation Report ⓘ

Oct 6, 2022 – 04:13 pm BST

PDB ID	:	7ZUS
Title	:	Crystal structure of ternary complex of Pol theta polymerase domain
Authors	:	Krajewski, W.W.; Turnbull, A.P.; Willis, S.; Charles, M.; Stockley, M.; Heald, R.A.
Deposited on	:	2022-05-13
Resolution	:	2.26 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

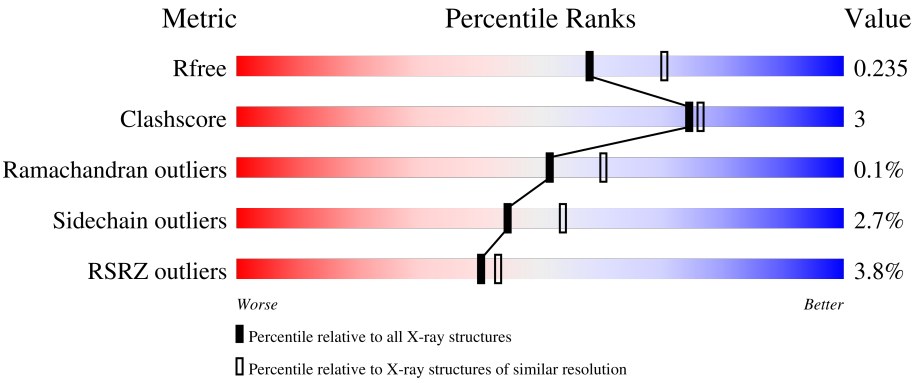
MolProbity	:	4.02b-467
Mogul	:	FAILED
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
buster-report	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

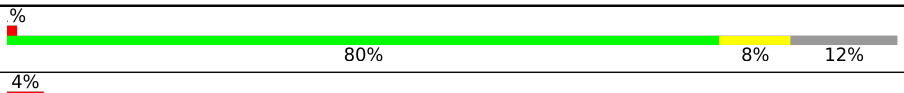
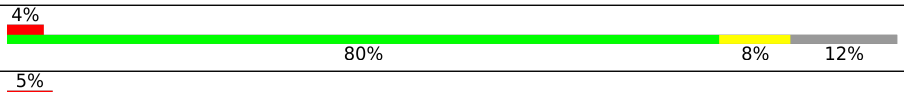
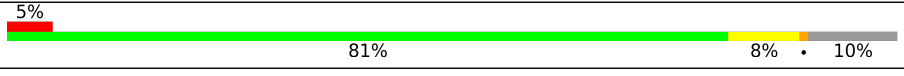
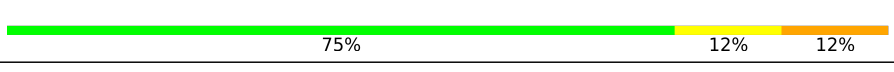
The reported resolution of this entry is 2.26 Å.

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	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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	AAA	726	
1	BBB	726	
1	CCC	726	
2	DDD	16	
2	FFF	16	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	HHH	16	
3	EEE	13	
3	GGG	13	
3	III	13	

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 16842 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 polymerase theta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	639	Total	C	N	O	S	0	0	0
			4882	3125	825	903	29			
1	BBB	639	Total	C	N	O	S	0	0	0
			4898	3135	828	905	30			
1	CCC	650	Total	C	N	O	S	0	0	0
			4937	3153	832	924	28			

There are 138 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	2261	GLY	PRO	engineered mutation	UNP O75417
AAA	?	-	THR	deletion	UNP O75417
AAA	?	-	LEU	deletion	UNP O75417
AAA	?	-	VAL	deletion	UNP O75417
AAA	?	-	GLY	deletion	UNP O75417
AAA	?	-	GLU	deletion	UNP O75417
AAA	?	-	SER	deletion	UNP O75417
AAA	?	-	PRO	deletion	UNP O75417
AAA	?	-	PRO	deletion	UNP O75417
AAA	?	-	SER	deletion	UNP O75417
AAA	?	-	GLN	deletion	UNP O75417
AAA	?	-	ALA	deletion	UNP O75417
AAA	?	-	VAL	deletion	UNP O75417
AAA	?	-	GLY	deletion	UNP O75417
AAA	?	-	LYS	deletion	UNP O75417
AAA	?	-	GLY	deletion	UNP O75417
AAA	?	-	LEU	deletion	UNP O75417
AAA	?	-	LEU	deletion	UNP O75417
AAA	?	-	PRO	deletion	UNP O75417
AAA	?	-	MET	deletion	UNP O75417
AAA	?	-	GLY	deletion	UNP O75417
AAA	?	-	ARG	deletion	UNP O75417
AAA	?	-	GLY	deletion	UNP O75417

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	?	-	LYS	deletion	UNP O75417
AAA	?	-	TYR	deletion	UNP O75417
AAA	?	-	LYS	deletion	UNP O75417
AAA	?	-	LYS	deletion	UNP O75417
AAA	?	-	GLY	deletion	UNP O75417
AAA	?	-	PHE	deletion	UNP O75417
AAA	?	-	SER	deletion	UNP O75417
AAA	?	-	VAL	deletion	UNP O75417
AAA	?	-	ASN	deletion	UNP O75417
AAA	?	-	PRO	deletion	UNP O75417
AAA	?	-	ARG	deletion	UNP O75417
AAA	?	-	CYS	deletion	UNP O75417
AAA	?	-	GLN	deletion	UNP O75417
AAA	?	-	ALA	deletion	UNP O75417
AAA	?	-	GLN	deletion	UNP O75417
AAA	?	-	MET	deletion	UNP O75417
AAA	?	-	GLU	deletion	UNP O75417
AAA	?	-	GLU	deletion	UNP O75417
AAA	?	-	ARG	deletion	UNP O75417
AAA	?	-	ALA	deletion	UNP O75417
AAA	?	-	ALA	deletion	UNP O75417
AAA	?	-	ASP	deletion	UNP O75417
AAA	?	-	ARG	deletion	UNP O75417
BBB	2261	GLY	PRO	engineered mutation	UNP O75417
BBB	?	-	THR	deletion	UNP O75417
BBB	?	-	LEU	deletion	UNP O75417
BBB	?	-	VAL	deletion	UNP O75417
BBB	?	-	GLY	deletion	UNP O75417
BBB	?	-	GLU	deletion	UNP O75417
BBB	?	-	SER	deletion	UNP O75417
BBB	?	-	PRO	deletion	UNP O75417
BBB	?	-	PRO	deletion	UNP O75417
BBB	?	-	SER	deletion	UNP O75417
BBB	?	-	GLN	deletion	UNP O75417
BBB	?	-	ALA	deletion	UNP O75417
BBB	?	-	VAL	deletion	UNP O75417
BBB	?	-	GLY	deletion	UNP O75417
BBB	?	-	LYS	deletion	UNP O75417
BBB	?	-	GLY	deletion	UNP O75417
BBB	?	-	LEU	deletion	UNP O75417
BBB	?	-	LEU	deletion	UNP O75417
BBB	?	-	PRO	deletion	UNP O75417

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	?	-	MET	deletion	UNP O75417
BBB	?	-	GLY	deletion	UNP O75417
BBB	?	-	ARG	deletion	UNP O75417
BBB	?	-	GLY	deletion	UNP O75417
BBB	?	-	LYS	deletion	UNP O75417
BBB	?	-	TYR	deletion	UNP O75417
BBB	?	-	LYS	deletion	UNP O75417
BBB	?	-	LYS	deletion	UNP O75417
BBB	?	-	GLY	deletion	UNP O75417
BBB	?	-	PHE	deletion	UNP O75417
BBB	?	-	SER	deletion	UNP O75417
BBB	?	-	VAL	deletion	UNP O75417
BBB	?	-	ASN	deletion	UNP O75417
BBB	?	-	PRO	deletion	UNP O75417
BBB	?	-	ARG	deletion	UNP O75417
BBB	?	-	CYS	deletion	UNP O75417
BBB	?	-	GLN	deletion	UNP O75417
BBB	?	-	ALA	deletion	UNP O75417
BBB	?	-	GLN	deletion	UNP O75417
BBB	?	-	MET	deletion	UNP O75417
BBB	?	-	GLU	deletion	UNP O75417
BBB	?	-	GLU	deletion	UNP O75417
BBB	?	-	ARG	deletion	UNP O75417
BBB	?	-	ALA	deletion	UNP O75417
BBB	?	-	ALA	deletion	UNP O75417
BBB	?	-	ASP	deletion	UNP O75417
BBB	?	-	ARG	deletion	UNP O75417
CCC	2261	GLY	PRO	engineered mutation	UNP O75417
CCC	?	-	THR	deletion	UNP O75417
CCC	?	-	LEU	deletion	UNP O75417
CCC	?	-	VAL	deletion	UNP O75417
CCC	?	-	GLY	deletion	UNP O75417
CCC	?	-	GLU	deletion	UNP O75417
CCC	?	-	SER	deletion	UNP O75417
CCC	?	-	PRO	deletion	UNP O75417
CCC	?	-	PRO	deletion	UNP O75417
CCC	?	-	SER	deletion	UNP O75417
CCC	?	-	GLN	deletion	UNP O75417
CCC	?	-	ALA	deletion	UNP O75417
CCC	?	-	VAL	deletion	UNP O75417
CCC	?	-	GLY	deletion	UNP O75417
CCC	?	-	LYS	deletion	UNP O75417

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
CCC	?	-	GLY	deletion	UNP O75417
CCC	?	-	LEU	deletion	UNP O75417
CCC	?	-	LEU	deletion	UNP O75417
CCC	?	-	PRO	deletion	UNP O75417
CCC	?	-	MET	deletion	UNP O75417
CCC	?	-	GLY	deletion	UNP O75417
CCC	?	-	ARG	deletion	UNP O75417
CCC	?	-	GLY	deletion	UNP O75417
CCC	?	-	LYS	deletion	UNP O75417
CCC	?	-	TYR	deletion	UNP O75417
CCC	?	-	LYS	deletion	UNP O75417
CCC	?	-	LYS	deletion	UNP O75417
CCC	?	-	GLY	deletion	UNP O75417
CCC	?	-	PHE	deletion	UNP O75417
CCC	?	-	SER	deletion	UNP O75417
CCC	?	-	VAL	deletion	UNP O75417
CCC	?	-	ASN	deletion	UNP O75417
CCC	?	-	PRO	deletion	UNP O75417
CCC	?	-	ARG	deletion	UNP O75417
CCC	?	-	CYS	deletion	UNP O75417
CCC	?	-	GLN	deletion	UNP O75417
CCC	?	-	ALA	deletion	UNP O75417
CCC	?	-	GLN	deletion	UNP O75417
CCC	?	-	MET	deletion	UNP O75417
CCC	?	-	GLU	deletion	UNP O75417
CCC	?	-	GLU	deletion	UNP O75417
CCC	?	-	ARG	deletion	UNP O75417
CCC	?	-	ALA	deletion	UNP O75417
CCC	?	-	ALA	deletion	UNP O75417
CCC	?	-	ASP	deletion	UNP O75417
CCC	?	-	ARG	deletion	UNP O75417

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	DDD	16	Total	C	N	O	P	0	0	0
			324	154	59	95	16			
2	FFF	16	Total	C	N	O	P	0	0	0
			324	154	59	95	16			
2	HHH	16	Total	C	N	O	P	0	0	0
			324	154	59	95	16			

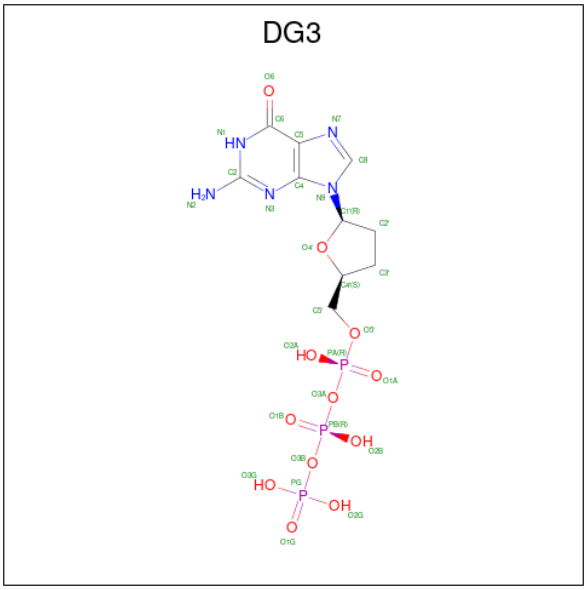
- Molecule 3 is a DNA chain called DNA (5'-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*TP*(DDG))-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	EEE	13	Total	C	N	O	P	0	0	0
			264	127	47	78	12			
3	GGG	13	Total	C	N	O	P	0	0	0
			264	127	47	78	12			
3	III	13	Total	C	N	O	P	0	0	0
			264	127	47	78	12			

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	1	Total	Mg	0	0
			1	1		
4	BBB	1	Total	Mg	0	0
			1	1		
4	CCC	1	Total	Mg	0	0
			1	1		

- Molecule 5 is 2'-3'-DIDEOXYGUANOSINE-5'-TRIPHOSPHATE (three-letter code: DG3) (formula: $C_{10}H_{16}N_5O_{12}P_3$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	AAA	1	Total	C	N	O	P	0	0
			30	10	5	12	3		
5	BBB	1	Total	C	N	O	P	0	0
			30	10	5	12	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	CCC	1	Total	C	N	O	P	0	0
			30	10	5	12	3		

- Molecule 6 is water.

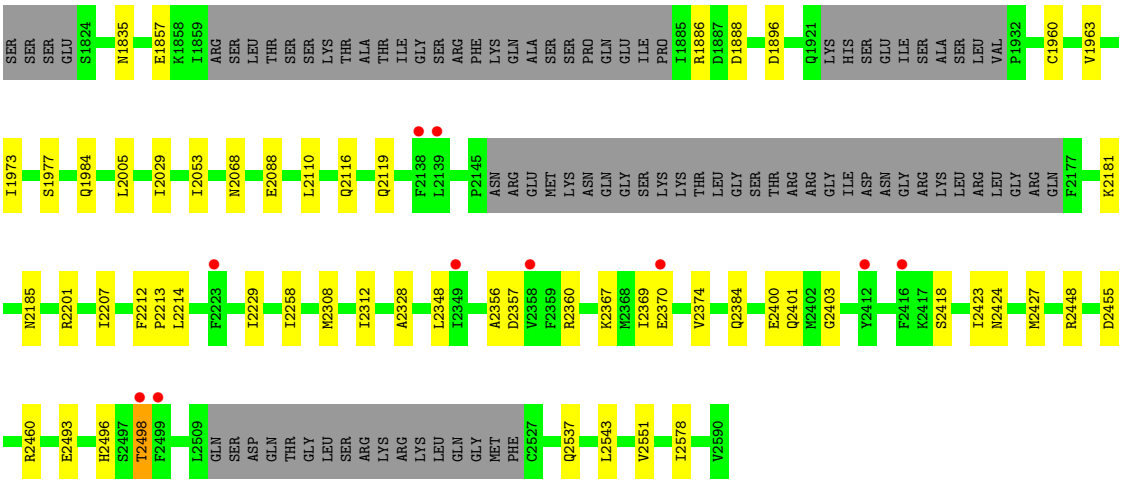
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	AAA	116	Total	O	0	0
			116	116		
6	BBB	71	Total	O	0	0
			71	71		
6	CCC	44	Total	O	0	0
			44	44		
6	DDD	6	Total	O	0	0
			6	6		
6	EEE	4	Total	O	0	0
			4	4		
6	FFF	4	Total	O	0	0
			4	4		
6	GGG	9	Total	O	0	0
			9	9		
6	HHH	10	Total	O	0	0
			10	10		
6	III	4	Total	O	0	0
			4	4		

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.

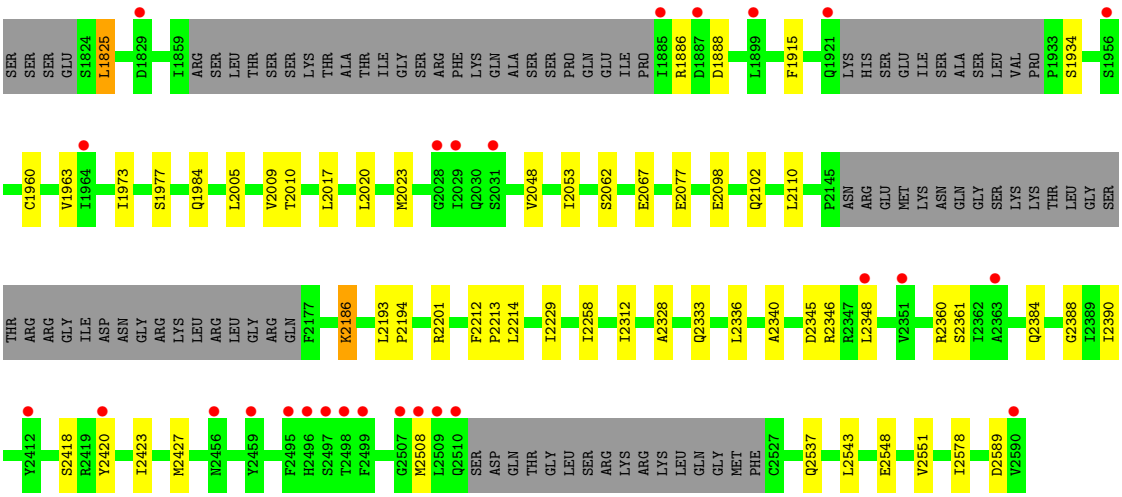
• Molecule 1: DNA polymerase theta

Chain AAA:



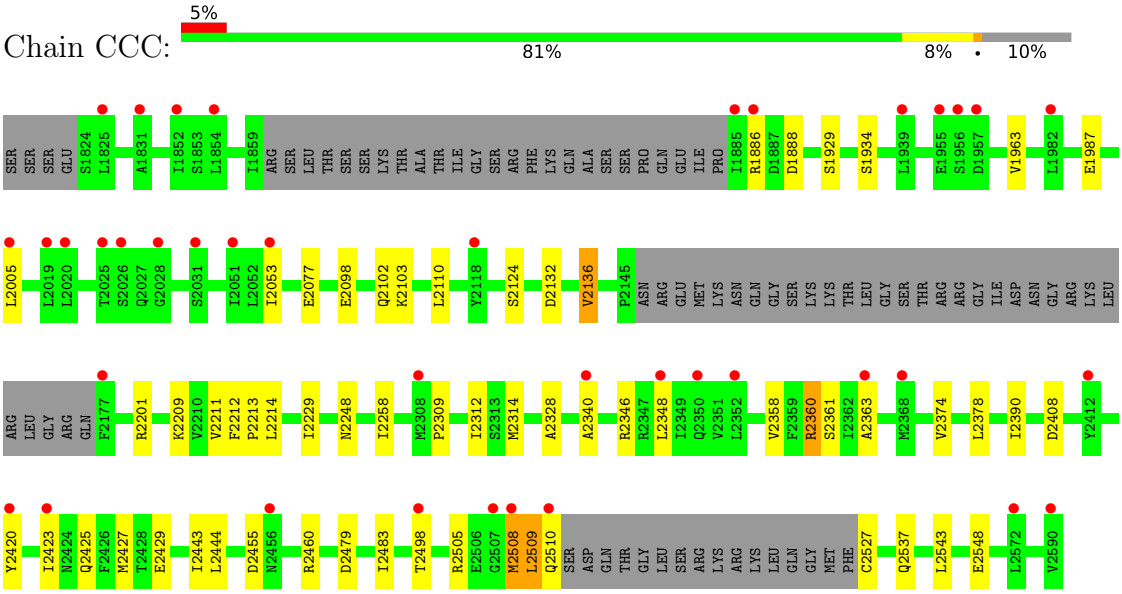
• Molecule 1: DNA polymerase theta

Chain BBB:



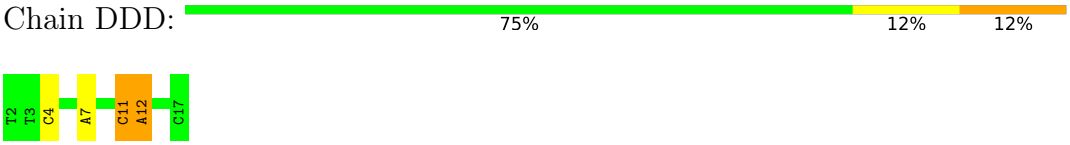
• Molecule 1: DNA polymerase theta

Chain CCC:



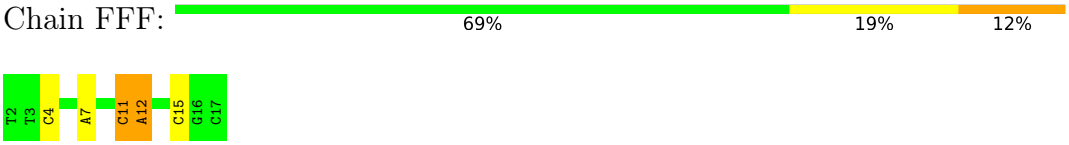
• Molecule 2: DNA (5’-D(P*TP*TP*CP*CP*AP*AP*TP*GP*AP*CP*AP*GP*CP*CP*GP*C)-3’)

Chain DDD:



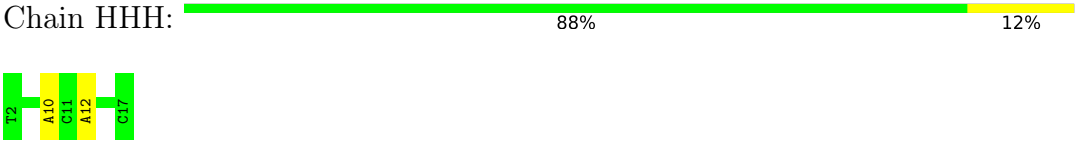
• Molecule 2: DNA (5’-D(P*TP*TP*CP*CP*AP*AP*TP*GP*AP*CP*AP*GP*CP*CP*GP*C)-3’)

Chain FFF:



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

Chain HHH:



• Molecule 3: DNA (5’-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*TP*(DDG))-3’)

Chain EEE:

• Molecule 3: DNA (5’-D(*GP*CP*GP*GP*CP*TP*GP*TP*CP*AP*TP*TP*(DDG))-3’)

Chain GGG:

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

Chain III:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	288.81Å 172.91Å 58.75Å 90.00° 90.89° 90.00°	Depositor
Resolution (Å)	50.01 – 2.26 84.10 – 2.26	Depositor EDS
% Data completeness (in resolution range)	98.9 (50.01-2.26) 98.9 (84.10-2.26)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.01 (at 2.27Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.201 , 0.238 0.202 , 0.235	Depositor DCC
R_{free} test set	6681 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	55.5	Xtriage
Anisotropy	0.452	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.009 for -1/2*h+3/2*k,1/2*h+1/2*k,-l 0.009 for -1/2*h-3/2*k,-1/2*h+1/2*k,-l 0.028 for 1/2*h+3/2*k,1/2*h-1/2*k,-l 0.023 for 1/2*h-3/2*k,-1/2*h-1/2*k,-l 0.024 for -h,-k,l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	16842	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.82% 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: DG3, DDG, 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	AAA	0.76	1/4978 (0.0%)	0.86	1/6748 (0.0%)
1	BBB	0.75	1/4994 (0.0%)	0.83	0/6766
1	CCC	0.71	1/5033 (0.0%)	0.81	0/6829
2	DDD	1.15	0/362	1.51	4/555 (0.7%)
2	FFF	1.17	0/362	1.45	4/555 (0.7%)
2	HHH	1.06	0/362	1.39	1/555 (0.2%)
3	EEE	1.26	2/271 (0.7%)	1.66	7/417 (1.7%)
3	GGG	1.36	3/271 (1.1%)	1.52	3/417 (0.7%)
3	III	1.22	1/271 (0.4%)	1.55	4/417 (1.0%)
All	All	0.80	9/16904 (0.1%)	0.95	24/23259 (0.1%)

All (9) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	GGG	11	DT	P-O5'	8.21	1.68	1.59
3	GGG	12	DT	P-O5'	8.03	1.67	1.59
3	III	10	DA	P-O5'	7.56	1.67	1.59
3	GGG	10	DA	P-O5'	6.60	1.66	1.59
3	EEE	11	DT	P-O5'	6.41	1.66	1.59
1	CCC	2077	GLU	CD-OE2	6.29	1.32	1.25
1	BBB	2077	GLU	CD-OE2	5.95	1.32	1.25
1	AAA	2088	GLU	CD-OE1	5.64	1.31	1.25
3	EEE	12	DT	P-O5'	5.17	1.65	1.59

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	III	7	DG	P-O3'-C3'	-12.77	104.38	119.70
3	GGG	7	DG	P-O3'-C3'	-12.10	105.18	119.70
3	EEE	7	DG	P-O3'-C3'	-11.36	106.06	119.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	EEE	9	DC	P-O3'-C3'	-10.10	107.58	119.70
3	EEE	12	DT	O5'-P-OP2	-8.31	98.22	105.70
2	DDD	12	DA	P-O3'-C3'	-6.94	111.37	119.70
2	DDD	7	DA	OP1-P-OP2	-6.84	109.34	119.60
2	FFF	12	DA	P-O3'-C3'	-6.70	111.66	119.70
1	AAA	2448	ARG	NE-CZ-NH2	-6.49	117.05	120.30
2	DDD	4	DC	O5'-P-OP2	-6.30	100.03	105.70
3	EEE	6	DT	P-O3'-C3'	-6.24	112.21	119.70
3	GGG	9	DC	OP2-P-O3'	6.09	118.60	105.20
3	III	9	DC	OP2-P-O3'	5.82	118.00	105.20
3	III	6	DT	P-O3'-C3'	-5.66	112.91	119.70
2	DDD	11	DC	O3'-P-O5'	-5.64	93.29	104.00
2	HHH	12	DA	P-O3'-C3'	-5.55	113.04	119.70
3	GGG	11	DT	OP2-P-O3'	5.49	117.27	105.20
2	FFF	7	DA	C4'-C3'-C2'	-5.33	98.30	103.10
3	EEE	11	DT	OP2-P-O3'	5.33	116.92	105.20
3	EEE	9	DC	OP2-P-O3'	5.30	116.86	105.20
3	EEE	10	DA	P-O3'-C3'	-5.30	113.34	119.70
2	FFF	15	DC	P-O3'-C3'	-5.24	113.41	119.70
2	FFF	11	DC	OP2-P-O3'	5.09	116.40	105.20
3	III	11	DT	OP2-P-O3'	5.08	116.38	105.20

There are no chirality outliers.

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	AAA	4882	0	4712	29	0
1	BBB	4898	0	4741	35	0
1	CCC	4937	0	4734	34	0
2	DDD	324	0	180	1	0
2	FFF	324	0	180	3	0
2	HHH	324	0	180	1	0
3	EEE	264	0	149	1	0
3	GGG	264	0	149	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	III	264	0	149	1	0
4	AAA	1	0	0	0	0
4	BBB	1	0	0	0	0
4	CCC	1	0	0	0	0
5	AAA	30	0	12	0	0
5	BBB	30	0	12	0	0
5	CCC	30	0	12	0	0
6	AAA	116	0	0	1	0
6	BBB	71	0	0	0	0
6	CCC	44	0	0	2	0
6	DDD	6	0	0	0	0
6	EEE	4	0	0	0	0
6	FFF	4	0	0	0	0
6	GGG	9	0	0	0	0
6	HHH	10	0	0	0	0
6	III	4	0	0	0	0
All	All	16842	0	15210	100	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 3.

All (100) 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:BBB:2340:ALA:HB2	1:BBB:2348:LEU:HD22	1.67	0.75
1:AAA:2496:HIS:CG	1:AAA:2498:THR:O	2.40	0.75
1:CCC:2390:ILE:O	1:CCC:2427:MET:HE1	1.86	0.74
1:CCC:2340:ALA:HB2	1:CCC:2348:LEU:HD22	1.69	0.74
1:BBB:2390:ILE:O	1:BBB:2427:MET:HE1	1.89	0.73
1:CCC:2348:LEU:HD12	1:CCC:2423:ILE:HD11	1.81	0.62
1:CCC:1963:VAL:CG2	1:CCC:2053:ILE:HG22	2.31	0.61
1:BBB:2348:LEU:HD12	1:BBB:2423:ILE:HD11	1.82	0.61
1:CCC:2505:ARG:NH2	1:CCC:2527:CYS:O	2.35	0.59
1:CCC:2360:ARG:NH2	1:CCC:2374:VAL:O	2.36	0.58
1:CCC:2425:GLN:O	1:CCC:2429:GLU:HG3	2.03	0.58
1:AAA:2357:ASP:HB3	1:AAA:2360:ARG:HB2	1.85	0.58
1:BBB:2009:VAL:HG12	1:BBB:2017:LEU:HD12	1.85	0.58
1:CCC:2005:LEU:HD13	1:CCC:2053:ILE:HD11	1.86	0.58
1:BBB:2340:ALA:HB2	1:BBB:2348:LEU:CD2	2.35	0.56
1:BBB:2258:ILE:HD12	1:BBB:2312:ILE:HG13	1.88	0.55
1:CCC:2340:ALA:HB2	1:CCC:2348:LEU:CD2	2.37	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:CCC:2098:GLU:O	1:CCC:2102:GLN:HG3	2.07	0.54
1:CCC:2132:ASP:O	1:CCC:2136:VAL:HG12	2.07	0.54
1:AAA:1857:GLU:OE2	6:AAA:2701:HOH:O	2.19	0.54
1:AAA:2214:LEU:HD22	1:AAA:2229:ILE:HD13	1.89	0.54
1:CCC:2505:ARG:O	1:CCC:2508:MET:HB3	2.09	0.53
1:CCC:2248:ASN:HB3	6:CCC:2720:HOH:O	2.08	0.53
1:AAA:2328:ALA:HB2	1:AAA:2543:LEU:CD2	2.39	0.52
1:BBB:2023:MET:CE	1:BBB:2048:VAL:HG11	2.39	0.52
2:DDD:11:DC:H2"	2:DDD:12:DA:C8	2.45	0.52
1:CCC:2328:ALA:HB2	1:CCC:2543:LEU:CD2	2.40	0.52
1:AAA:2360:ARG:HG2	1:AAA:2374:VAL:HG21	1.92	0.52
1:BBB:2328:ALA:HB2	1:BBB:2543:LEU:CD2	2.40	0.52
1:AAA:2367:LYS:O	1:AAA:2369:ILE:HG13	2.11	0.51
1:BBB:2020:LEU:HA	1:BBB:2023:MET:HE2	1.92	0.51
1:AAA:2374:VAL:O	1:AAA:2374:VAL:HG23	2.12	0.50
1:AAA:2493:GLU:HA	1:AAA:2496:HIS:HB2	1.92	0.50
1:AAA:2110:LEU:HD12	1:AAA:2207:ILE:HD12	1.92	0.50
3:EEE:9:DC:C2'	3:EEE:10:DA:OP2	2.60	0.50
1:BBB:2348:LEU:C	1:BBB:2348:LEU:HD23	2.33	0.49
1:AAA:2308:MET:HE2	1:BBB:2186:LYS:HG2	1.95	0.49
1:BBB:2420:TYR:HB3	1:BBB:2423:ILE:HD12	1.95	0.49
1:CCC:2348:LEU:HD23	1:CCC:2348:LEU:C	2.34	0.47
1:CCC:2214:LEU:HD22	1:CCC:2229:ILE:HD13	1.96	0.47
1:AAA:2116:GLN:HA	1:AAA:2119:GLN:HG2	1.97	0.47
1:BBB:2214:LEU:HD22	1:BBB:2229:ILE:HD13	1.96	0.46
1:BBB:2010:THR:HA	1:BBB:2017:LEU:HD11	1.97	0.46
1:CCC:1963:VAL:HG23	1:CCC:2053:ILE:HG22	1.97	0.46
1:CCC:2209:LYS:HE3	2:HHH:10:DA:H4'	1.97	0.46
1:AAA:2455:ASP:OD2	1:AAA:2460:ARG:HD2	2.15	0.46
1:BBB:2384:GLN:HG3	2:FFF:4:DC:N4	2.30	0.46
1:AAA:2005:LEU:HD13	1:AAA:2053:ILE:HD11	1.97	0.46
1:BBB:2098:GLU:O	1:BBB:2102:GLN:HG3	2.15	0.46
1:CCC:2258:ILE:HD12	1:CCC:2312:ILE:HG13	1.98	0.46
1:BBB:2005:LEU:HD13	1:BBB:2053:ILE:HD11	1.98	0.46
1:BBB:2551:VAL:HB	1:BBB:2578:ILE:HD11	1.98	0.45
1:BBB:2023:MET:HE3	1:BBB:2048:VAL:HG11	1.97	0.45
1:CCC:2103:LYS:HD2	1:CCC:2211:VAL:HG11	1.99	0.45
1:CCC:2110:LEU:HD12	1:CCC:2110:LEU:HA	1.79	0.45
1:CCC:2314:MET:HG3	6:CCC:2735:HOH:O	2.16	0.45
1:CCC:2443:ILE:HD11	1:CCC:2479:ASP:HB2	1.98	0.45
1:AAA:2348:LEU:HD22	1:AAA:2423:ILE:HD11	1.98	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:CCC:2455:ASP:OD2	1:CCC:2460:ARG:HD2	2.16	0.45
1:CCC:2212:PHE:HB3	1:CCC:2213:PRO:HD3	1.99	0.44
1:CCC:1886:ARG:NH2	1:CCC:1888:ASP:OD1	2.49	0.44
1:CCC:2420:TYR:HB3	1:CCC:2423:ILE:HD12	2.00	0.44
2:FFF:11:DC:H2"	2:FFF:12:DA:C8	2.53	0.44
1:AAA:1960:CYS:HB3	1:AAA:1984:GLN:HG2	2.00	0.44
1:BBB:1886:ARG:NH2	1:BBB:1888:ASP:OD1	2.50	0.44
1:AAA:2212:PHE:HB3	1:AAA:2213:PRO:HD3	2.00	0.44
1:AAA:2214:LEU:HD22	1:AAA:2229:ILE:CD1	2.48	0.43
1:BBB:2333:GLN:HB3	1:BBB:2336:LEU:HB2	1.99	0.43
1:BBB:1825:LEU:HD12	1:BBB:1825:LEU:C	2.38	0.43
1:AAA:2110:LEU:CD1	1:AAA:2207:ILE:HD12	2.48	0.43
1:BBB:2110:LEU:HD12	1:BBB:2110:LEU:HA	1.82	0.43
1:BBB:2388:GLY:HA2	2:FFF:4:DC:C2	2.54	0.43
1:AAA:2424:ASN:HA	1:AAA:2427:MET:HE2	2.00	0.42
1:BBB:2023:MET:HE1	1:BBB:2048:VAL:HG11	2.01	0.42
1:CCC:2509:LEU:O	1:CCC:2510:GLN:HG2	2.20	0.42
1:BBB:2212:PHE:HB3	1:BBB:2213:PRO:HD3	2.00	0.42
1:BBB:2508:MET:HE2	1:BBB:2508:MET:HB3	1.91	0.42
1:AAA:2258:ILE:HD12	1:AAA:2312:ILE:HG13	2.02	0.42
1:AAA:2493:GLU:HA	1:AAA:2496:HIS:CB	2.49	0.42
1:CCC:2348:LEU:HD23	1:CCC:2348:LEU:O	2.18	0.42
1:BBB:1963:VAL:HG12	1:BBB:2053:ILE:HG22	2.01	0.42
1:AAA:1886:ARG:NH2	1:AAA:1888:ASP:OD1	2.52	0.42
1:AAA:2369:ILE:HG22	1:AAA:2370:GLU:N	2.35	0.41
1:BBB:1973:ILE:O	1:BBB:1977:SER:HB2	2.19	0.41
1:AAA:2493:GLU:O	1:AAA:2496:HIS:HB3	2.20	0.41
1:AAA:2551:VAL:HB	1:AAA:2578:ILE:HD11	2.01	0.41
1:BBB:1960:CYS:HB3	1:BBB:1984:GLN:HG2	2.01	0.41
1:CCC:2363:ALA:HB1	1:CCC:2378:LEU:HG	2.02	0.41
3:III:5:DC:H2"	3:III:6:DT:OP2	2.21	0.41
1:AAA:1973:ILE:O	1:AAA:1977:SER:HB2	2.21	0.41
1:CCC:2444:LEU:HD21	1:CCC:2483:ILE:HD11	2.03	0.41
1:BBB:2348:LEU:HD23	1:BBB:2348:LEU:O	2.20	0.41
1:BBB:2193:LEU:HB3	1:BBB:2194:PRO:HD3	2.01	0.40
1:CCC:1963:VAL:HA	1:CCC:1987:GLU:O	2.21	0.40
1:AAA:1963:VAL:HG12	1:AAA:2053:ILE:HG22	2.03	0.40
1:AAA:2403:GLY:O	1:CCC:2309:PRO:HD3	2.22	0.40
1:BBB:2214:LEU:HD22	1:BBB:2229:ILE:CD1	2.52	0.40
1:BBB:2345:ASP:OD2	1:BBB:2420:TYR:HD1	2.03	0.40
1:CCC:2214:LEU:HD22	1:CCC:2229:ILE:CD1	2.51	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:BBB:2390:ILE:O	1:BBB:2427:MET:CE	2.66	0.40

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	AAA	629/726 (87%)	612 (97%)	16 (2%)	1 (0%)	47 55
1	BBB	629/726 (87%)	615 (98%)	14 (2%)	0	100 100
1	CCC	642/726 (88%)	627 (98%)	14 (2%)	1 (0%)	47 55
All	All	1900/2178 (87%)	1854 (98%)	44 (2%)	2 (0%)	51 60

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	2356	ALA
1	CCC	2498	THR

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	AAA	499/638 (78%)	486 (97%)	13 (3%)	46 55
1	BBB	502/638 (79%)	488 (97%)	14 (3%)	43 52

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	CCC	503/638 (79%)	489 (97%)	14 (3%)	43 52
All	All	1504/1914 (79%)	1463 (97%)	41 (3%)	44 54

All (41) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	1835	ASN
1	AAA	1896	ASP
1	AAA	2029	ILE
1	AAA	2068	ASN
1	AAA	2181	LYS
1	AAA	2185	ASN
1	AAA	2201	ARG
1	AAA	2384	GLN
1	AAA	2400	GLU
1	AAA	2401	GLN
1	AAA	2418	SER
1	AAA	2498	THR
1	AAA	2537	GLN
1	BBB	1825	LEU
1	BBB	1915	PHE
1	BBB	1934	SER
1	BBB	2062	SER
1	BBB	2067	GLU
1	BBB	2186	LYS
1	BBB	2201	ARG
1	BBB	2346	ARG
1	BBB	2360	ARG
1	BBB	2361	SER
1	BBB	2418	SER
1	BBB	2537	GLN
1	BBB	2548	GLU
1	BBB	2589	ASP
1	CCC	1929	SER
1	CCC	1934	SER
1	CCC	2124	SER
1	CCC	2136	VAL
1	CCC	2201	ARG
1	CCC	2346	ARG
1	CCC	2358	VAL
1	CCC	2360	ARG
1	CCC	2361	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	CCC	2408	ASP
1	CCC	2508	MET
1	CCC	2509	LEU
1	CCC	2537	GLN
1	CCC	2548	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

Mogul failed to run properly - this section is therefore empty.

5.5 Carbohydrates (i)

Mogul failed to run properly - this section is therefore empty.

5.6 Ligand geometry (i)

Mogul failed to run properly - this section is therefore empty.

5.7 Other polymers (i)

Mogul failed to run properly - this section is therefore empty.

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	AAA	639/726 (88%)	0.28	10 (1%) 72 74	38, 65, 113, 142	0
1	BBB	639/726 (88%)	0.31	27 (4%) 36 38	44, 72, 112, 132	0
1	CCC	650/726 (89%)	0.39	39 (6%) 21 23	48, 84, 121, 143	0
2	DDD	16/16 (100%)	0.05	0 100 100	49, 99, 173, 183	0
2	FFF	16/16 (100%)	-0.14	0 100 100	48, 80, 157, 165	0
2	HHH	16/16 (100%)	-0.41	0 100 100	55, 92, 147, 151	0
3	EEE	12/13 (92%)	-0.04	0 100 100	50, 110, 155, 160	0
3	GGG	12/13 (92%)	-0.35	0 100 100	50, 102, 142, 146	0
3	III	12/13 (92%)	0.04	0 100 100	53, 105, 147, 151	0
All	All	2012/2265 (88%)	0.30	76 (3%) 40 43	38, 75, 119, 183	0

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	2363	ALA	5.0
1	CCC	2510	GLN	4.7
1	BBB	2029	ILE	4.5
1	CCC	2590	VAL	4.2
1	AAA	2416	PHE	4.2
1	BBB	1885	ILE	3.7
1	AAA	2412	TYR	3.7
1	CCC	2352	LEU	3.5
1	BBB	2351	VAL	3.5
1	BBB	2507	GLY	3.4
1	CCC	1956	SER	3.3
1	CCC	2423	ILE	3.3
1	CCC	2028	GLY	3.3
1	BBB	2348	LEU	3.3
1	CCC	2363	ALA	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	CCC	2177	PHE	3.2
1	CCC	2118	TYR	3.0
1	BBB	2590	VAL	3.0
1	CCC	2348	LEU	3.0
1	CCC	2507	GLY	3.0
1	CCC	1982	LEU	2.9
1	CCC	2020	LEU	2.9
1	CCC	2572	LEU	2.9
1	BBB	1921	GLN	2.9
1	AAA	2139	LEU	2.8
1	BBB	2420	TYR	2.8
1	CCC	1852	ILE	2.8
1	CCC	2026	SER	2.8
1	CCC	2498	THR	2.8
1	BBB	2499	PHE	2.7
1	AAA	2370	GLU	2.7
1	CCC	2051	ILE	2.7
1	CCC	2031	SER	2.7
1	CCC	2019	LEU	2.7
1	BBB	2028	GLY	2.7
1	BBB	2497	SER	2.6
1	BBB	2508	MET	2.6
1	BBB	2031	SER	2.6
1	CCC	1831	ALA	2.6
1	CCC	2025	THR	2.5
1	BBB	2498	THR	2.5
1	CCC	1939	LEU	2.5
1	CCC	2508	MET	2.5
1	BBB	2412	TYR	2.5
1	CCC	2456	ASN	2.5
1	BBB	2509	LEU	2.4
1	CCC	1957	ASP	2.4
1	CCC	2340	ALA	2.4
1	BBB	2510	GLN	2.4
1	BBB	1956	SER	2.4
1	CCC	1825	LEU	2.3
1	CCC	2368	MET	2.3
1	BBB	2496	HIS	2.3
1	AAA	2358	VAL	2.3
1	BBB	2495	PHE	2.3
1	CCC	2350	GLN	2.3
1	AAA	2499	PHE	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	CCC	2412	TYR	2.2
1	CCC	1886	ARG	2.2
1	CCC	1885	ILE	2.2
1	CCC	1955	GLU	2.2
1	AAA	2498	THR	2.2
1	BBB	1964	ILE	2.1
1	CCC	2005	LEU	2.1
1	AAA	2138	PHE	2.1
1	CCC	1854	LEU	2.1
1	BBB	1829	ASP	2.1
1	BBB	1887	ASP	2.1
1	BBB	2459	TYR	2.1
1	BBB	1899	LEU	2.0
1	CCC	2053	ILE	2.0
1	AAA	2223	PHE	2.0
1	BBB	2456	ASN	2.0
1	CCC	2420	TYR	2.0
1	AAA	2349	ILE	2.0
1	CCC	2308	MET	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (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
3	DDG	EEE	13	21/22	0.97	0.16	45,48,53,56	0
3	DDG	GGG	13	21/22	0.97	0.14	45,49,55,58	0
3	DDG	III	13	21/22	0.97	0.13	50,52,59,63	0

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
4	MG	CCC	2601	1/1	0.85	0.12	63,63,63,63	0
4	MG	AAA	2601	1/1	0.89	0.05	55,55,55,55	0
4	MG	BBB	2601	1/1	0.96	0.13	64,64,64,64	0
5	DG3	CCC	2602	30/30	0.97	0.13	48,55,65,68	0
5	DG3	BBB	2602	30/30	0.98	0.13	48,56,63,68	0
5	DG3	AAA	2602	30/30	0.98	0.15	41,48,60,65	0

6.5 Other polymers (i)

There are no such residues in this entry.