wwPDB X-ray Structure Validation Summary Report (i)

May 15, 2020 – 12:52 pm BST

PDB ID	:	4TRT
Title	:	Deinococcus radiodurans DNA polymerase III subunit beta
Authors	:	Niiranen, L.; Lian, K.; Johnson, K.A.; Moe, E.
Deposited on	:	2014-06-17
Resolution	:	2.00 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 following versions of software and data (see references (1)) were used in the production of this report:

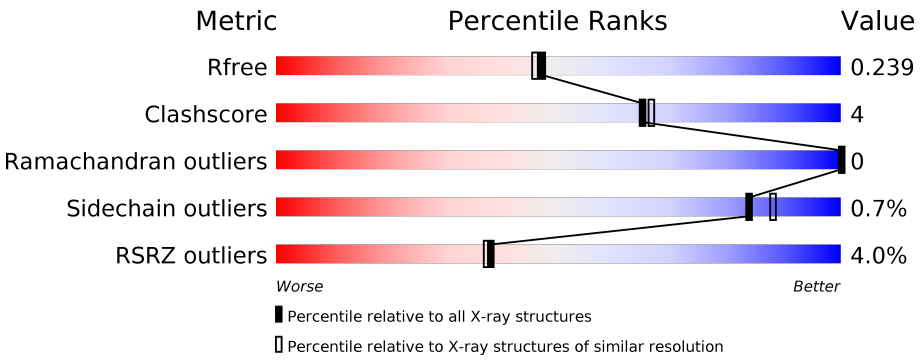
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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 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	368	4% 89% 9% •
1	B	368	4% 90% 8% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5758 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 III subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	361	2739	1715	480	536	8	0	3	0
1	B	361	2727	1709	478	533	7	0	1	0

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	VAL	MET	conflict	UNP Q9RYE8
A	347	VAL	ARG	see remark 999	UNP Q9RYE8
A	348	GLY	ARG	see remark 999	UNP Q9RYE8
A	349	GLY	TRP	see remark 999	UNP Q9RYE8
A	351	GLY	ARG	see remark 999	UNP Q9RYE8
A	352	GLY	ARG	see remark 999	UNP Q9RYE8
A	353	TYR	VAL	see remark 999	UNP Q9RYE8
A	354	MET	TYR	see remark 999	UNP Q9RYE8
A	355	ALA	GLY	see remark 999	UNP Q9RYE8
A	356	VAL	GLY	see remark 999	UNP Q9RYE8
A	357	MET	HIS	see remark 999	UNP Q9RYE8
A	358	VAL	GLY	see remark 999	UNP Q9RYE8
A	359	THR	HIS	see remark 999	UNP Q9RYE8
A	360	LEU	ALA	see remark 999	UNP Q9RYE8
A	361	ARG	ALA	see remark 999	UNP Q9RYE8
A	362	VAL	-	see remark 999	UNP Q9RYE8
A	363	GLU	-	expression tag	UNP Q9RYE8
A	364	ASN	-	expression tag	UNP Q9RYE8
A	365	LEU	-	expression tag	UNP Q9RYE8
A	366	TYR	-	expression tag	UNP Q9RYE8
A	367	PHE	-	expression tag	UNP Q9RYE8
A	368	GLN	-	expression tag	UNP Q9RYE8
B	1	VAL	MET	conflict	UNP Q9RYE8
B	347	VAL	ARG	see remark 999	UNP Q9RYE8
B	348	GLY	ARG	see remark 999	UNP Q9RYE8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	349	GLY	TRP	see remark 999	UNP Q9RYE8
B	351	GLY	ARG	see remark 999	UNP Q9RYE8
B	352	GLY	ARG	see remark 999	UNP Q9RYE8
B	353	TYR	VAL	see remark 999	UNP Q9RYE8
B	354	MET	TYR	see remark 999	UNP Q9RYE8
B	355	ALA	GLY	see remark 999	UNP Q9RYE8
B	356	VAL	GLY	see remark 999	UNP Q9RYE8
B	357	MET	HIS	see remark 999	UNP Q9RYE8
B	358	VAL	GLY	see remark 999	UNP Q9RYE8
B	359	THR	HIS	see remark 999	UNP Q9RYE8
B	360	LEU	ALA	see remark 999	UNP Q9RYE8
B	361	ARG	ALA	see remark 999	UNP Q9RYE8
B	362	VAL	-	see remark 999	UNP Q9RYE8
B	363	GLU	-	expression tag	UNP Q9RYE8
B	364	ASN	-	expression tag	UNP Q9RYE8
B	365	LEU	-	expression tag	UNP Q9RYE8
B	366	TYR	-	expression tag	UNP Q9RYE8
B	367	PHE	-	expression tag	UNP Q9RYE8
B	368	GLN	-	expression tag	UNP Q9RYE8

- Molecule 2 is water.

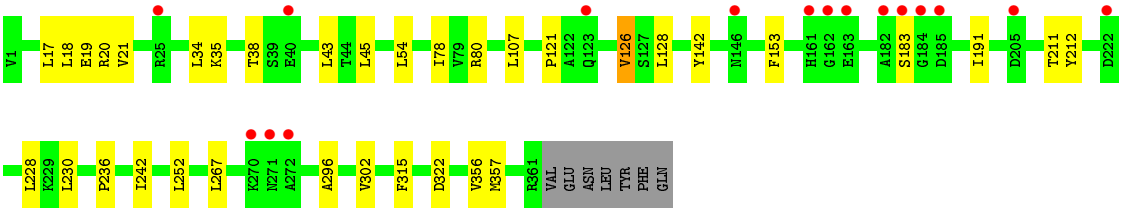
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	147	147	147	0	0
			Total	O		
2	B	145	145	145	0	0

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.

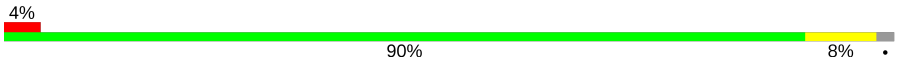
- Molecule 1: DNA polymerase III subunit beta

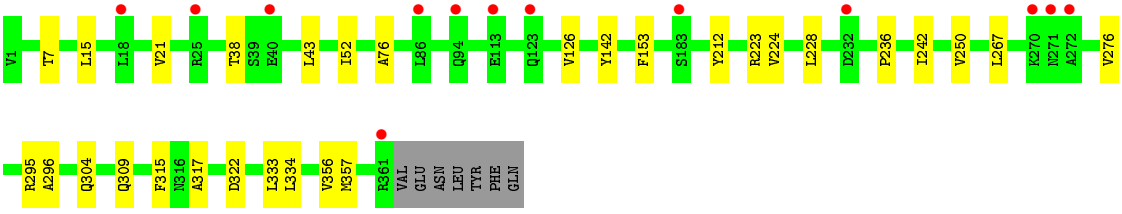
Chain A:



- Molecule 1: DNA polymerase III subunit beta

Chain B:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	84.41Å 84.41Å 198.74Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.00 28.94 – 2.00	Depositor EDS
% Data completeness (in resolution range)	98.0 (30.00-2.00) 98.1 (28.94-2.00)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.75 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.198 , 0.235 0.201 , 0.239	Depositor DCC
R_{free} test set	2806 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	32.3	Xtriage
Anisotropy	0.050	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 48.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	0.032 for -h,-k,l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	5758	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% 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 (i)

5.1 Standard geometry (i)

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	0.44	0/2779	0.66	0/3762
1	B	0.42	0/2767	0.63	0/3746
All	All	0.43	0/5546	0.65	0/7508

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	A	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	183	SER	Peptide

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	A	2739	0	2751	23	0
1	B	2727	0	2743	22	0
2	A	147	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	145	0	0	2	0
All	All	5758	0	5494	45	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 4.

The worst 5 of 45 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:191:ILE:HD13	1:A:230:LEU:CD2	2.13	0.79
1:A:38:THR:HG22	1:A:43:LEU:HD23	1.73	0.70
1:B:126:VAL:HG21	1:B:212:TYR:OH	1.93	0.67
1:B:267:LEU:HD12	1:B:296:ALA:HB2	1.83	0.61
1:A:35:LYS:NZ	2:A:478:HOH:O	2.34	0.59

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	362/368 (98%)	351 (97%)	11 (3%)	0	100	100
1	B	360/368 (98%)	354 (98%)	6 (2%)	0	100	100
All	All	722/736 (98%)	705 (98%)	17 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	293/297 (99%)	290 (99%)	3 (1%)	76 81
1	B	291/297 (98%)	290 (100%)	1 (0%)	92 95
All	All	584/594 (98%)	580 (99%)	4 (1%)	84 88

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	126	VAL
1	A	128	LEU
1	A	228	LEU
1	B	228	LEU

Some 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)

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

5.5 Carbohydrates (i)

There are no carbohydrates 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)

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	361/368 (98%)	0.14	16 (4%)	34	33	31, 42, 67, 93	0
1	B	361/368 (98%)	0.15	13 (3%)	42	42	29, 44, 66, 82	0
All	All	722/736 (98%)	0.14	29 (4%)	38	37	29, 43, 67, 93	0

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	183	SER	5.8
1	A	270	LYS	5.8
1	A	272	ALA	5.5
1	A	163	GLU	5.2
1	A	184	GLY	5.1

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 carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.