



Full wwPDB X-ray Structure Validation Report ⓘ

Dec 22, 2020 – 02:04 pm GMT

PDB ID : 7B06
Title : TgoT_RT521 apo
Authors : Samson, C.; Legrand, P.; Tekpinar, M.; Rozenski, J.; Abramov, M.; Holliger, P.; Pinheiro, V.; Herdewijn, P.; Delarue, M.
Deposited on : 2020-11-18
Resolution : 2.35 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.16
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) : 2.16

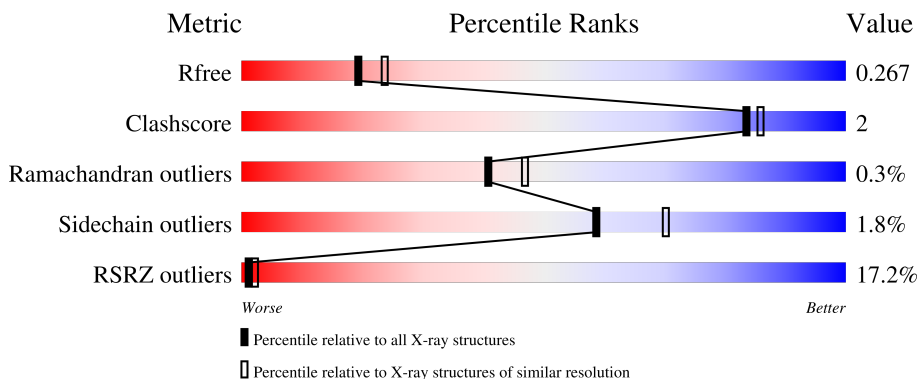
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.35 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	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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 $\leq 5\%$. 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	773	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5759 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.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	685	5641	3641	937	1050	13	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	93	GLN	VAL	conflict	UNP P56689
A	141	ALA	ASP	conflict	UNP P56689
A	143	ALA	GLU	conflict	UNP P56689
A	485	LEU	ALA	conflict	UNP P56689
A	521	LEU	ILE	conflict	UNP P56689
A	664	LYS	GLU	conflict	UNP P56689

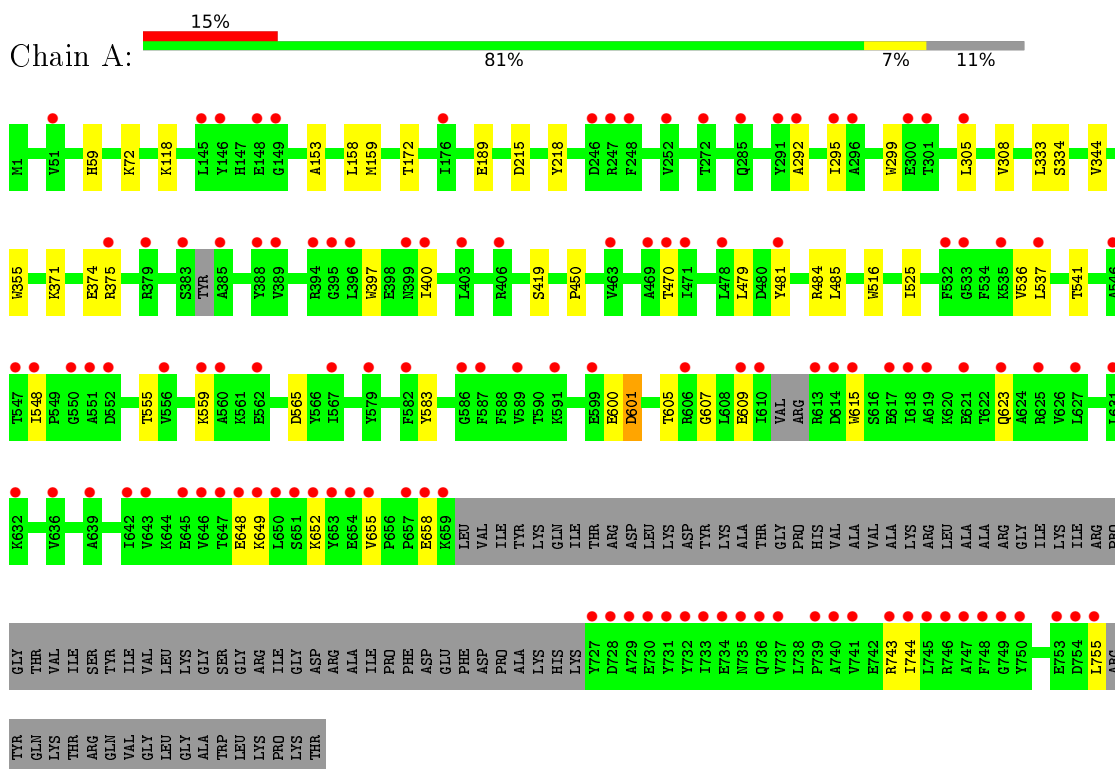
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	118	Total	O	0	0
			118	118		

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



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	70.53Å 109.85Å 111.30Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.69 – 2.35 43.69 – 2.35	Depositor EDS
% Data completeness (in resolution range)	99.3 (43.69-2.35) 99.2 (43.69-2.35)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.29 (at 2.34Å)	Xtrriage
Refinement program	BUSTER 2.10.3 (20-MAY-2020)	Depositor
R, R_{free}	0.225 , 0.247 0.247 , 0.267	Depositor DCC
R_{free} test set	1825 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	56.6	Xtrriage
Anisotropy	0.522	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 51.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.003 for -h,l,k	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	5759	wwPDB-VP
Average B, all atoms (Å ²)	95.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.00% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical 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 [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	0.28	0/5766	0.47	0/7772

There are no bond length outliers.

There are no bond angle outliers.

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	A	5641	0	5639	27	0
2	A	118	0	0	0	0
All	All	5759	0	5639	27	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 2.

All (27) 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:158:LEU:HD22	1:A:299:TRP:CD2	2.20	0.76
1:A:555:THR:HG22	1:A:559:LYS:HE2	1.77	0.65
1:A:623:GLN:HE21	1:A:744:ILE:HD11	1.63	0.64
1:A:158:LEU:HD22	1:A:299:TRP:CE2	2.33	0.64

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:548:ILE:HG13	1:A:559:LYS:HE3	1.82	0.60
1:A:118:LYS:HZ3	1:A:355:TRP:HZ2	1.50	0.60
1:A:555:THR:O	1:A:559:LYS:HG2	2.01	0.59
1:A:609:GLU:H	1:A:623:GLN:HE22	1.51	0.57
1:A:605:THR:HG22	1:A:607:GLY:H	1.71	0.56
1:A:400:ILE:HG13	1:A:583:TYR:HB2	1.88	0.54
1:A:648:GLU:HG2	1:A:652:LYS:HE2	1.90	0.54
1:A:305:LEU:HA	1:A:308:VAL:HG12	1.89	0.53
1:A:419:SER:HB3	1:A:450:PRO:HD3	1.92	0.51
1:A:397:TRP:CE2	1:A:537:LEU:HD22	2.47	0.50
1:A:159:MET:HG2	1:A:172:THR:HB	1.95	0.48
1:A:525:ILE:HG23	1:A:536:VAL:HG21	1.97	0.46
1:A:481:TYR:HD1	1:A:484:ARG:HH21	1.62	0.45
1:A:333:LEU:HD13	1:A:485:LEU:HD22	1.99	0.44
1:A:615:TRP:HH2	1:A:743:ARG:HH22	1.65	0.44
1:A:470:THR:HG21	1:A:479:LEU:HD12	2.00	0.43
1:A:118:LYS:NZ	1:A:355:TRP:HZ2	2.17	0.42
1:A:334:SER:HA	1:A:344:VAL:HG21	2.01	0.42
1:A:649:LYS:HB3	1:A:655:VAL:HG23	2.02	0.41
1:A:153:ALA:HA	1:A:218:TYR:CZ	2.56	0.41
1:A:292:ALA:O	1:A:295:ILE:HG12	2.20	0.40
1:A:600:GLU:O	1:A:601:ASP:HB2	2.21	0.40
1:A:292:ALA:HA	1:A:295:ILE:HG12	2.04	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	A	677/773 (88%)	655 (97%)	20 (3%)	2 (0%)	41 47

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	601	ASP
1	A	541	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	A	596/669 (89%)	585 (98%)	11 (2%)	59 70

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	59	HIS
1	A	72	LYS
1	A	189	GLU
1	A	215	ASP
1	A	371	LYS
1	A	374	GLU
1	A	375	ARG
1	A	516	TRP
1	A	565	ASP
1	A	658	GLU
1	A	755	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	91	GLN
1	A	623	GLN

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 monosaccharides 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

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	685/773 (88%)	1.22	118 (17%) 1 2	55, 86, 173, 190	0

All (118) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	727	TYR	15.8
1	A	653	TYR	15.7
1	A	613	ARG	15.3
1	A	750	TYR	13.5
1	A	737	VAL	13.3
1	A	650	LEU	9.9
1	A	729	ALA	9.8
1	A	395	GLY	8.8
1	A	732	TYR	8.7
1	A	747	ALA	8.6
1	A	659	LYS	8.5
1	A	639	ALA	8.3
1	A	731	TYR	8.2
1	A	647	THR	7.8
1	A	625	ARG	7.6
1	A	610	ILE	7.4
1	A	753	GLU	7.2
1	A	728	ASP	7.0
1	A	749	GLY	6.9
1	A	655	VAL	6.8
1	A	643	VAL	6.5
1	A	615	TRP	6.4
1	A	657	PRO	6.0
1	A	645	GLU	5.8
1	A	733	ILE	5.2
1	A	636	VAL	5.1
1	A	631	LEU	5.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	471	ILE	5.0
1	A	755	LEU	4.9
1	A	654	GLU	4.8
1	A	651	SER	4.8
1	A	560	ALA	4.8
1	A	385	ALA	4.7
1	A	743	ARG	4.6
1	A	301	THR	4.5
1	A	589	VAL	4.5
1	A	646	VAL	4.5
1	A	744	ILE	4.3
1	A	295	ILE	4.2
1	A	652	LYS	4.2
1	A	741	VAL	4.2
1	A	548	ILE	4.2
1	A	579	TYR	4.1
1	A	176	ILE	4.0
1	A	649	LYS	4.0
1	A	587	PHE	4.0
1	A	735	ASN	4.0
1	A	552	ASP	4.0
1	A	470	THR	3.9
1	A	146	TYR	3.8
1	A	292	ALA	3.8
1	A	609	GLU	3.7
1	A	396	LEU	3.7
1	A	478	LEU	3.7
1	A	375	ARG	3.6
1	A	532	PHE	3.5
1	A	559	LYS	3.5
1	A	400	ILE	3.5
1	A	627	LEU	3.5
1	A	739	PRO	3.5
1	A	383	SER	3.4
1	A	248	PHE	3.4
1	A	379	ARG	3.4
1	A	730	GLU	3.3
1	A	550	GLY	3.3
1	A	537	LEU	3.3
1	A	748	PHE	3.2
1	A	599	GLU	3.2
1	A	736	GLN	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	246	ASP	3.1
1	A	547	THR	3.0
1	A	582	PHE	3.0
1	A	619	ALA	3.0
1	A	617	GLU	3.0
1	A	614	ASP	2.9
1	A	734	GLU	2.9
1	A	618	ILE	2.8
1	A	533	GLY	2.8
1	A	148	GLU	2.8
1	A	394	ARG	2.8
1	A	535	LYS	2.8
1	A	463	VAL	2.7
1	A	746	ARG	2.7
1	A	658	GLU	2.7
1	A	389	VAL	2.6
1	A	388	TYR	2.6
1	A	562	GLU	2.6
1	A	403	LEU	2.5
1	A	606	ARG	2.5
1	A	546	ALA	2.5
1	A	556	VAL	2.5
1	A	272	THR	2.5
1	A	51	VAL	2.4
1	A	247	ARG	2.4
1	A	623	GLN	2.4
1	A	648	GLU	2.4
1	A	291	TYR	2.4
1	A	406	ARG	2.3
1	A	621	GLU	2.2
1	A	591	LYS	2.2
1	A	551	ALA	2.2
1	A	740	ALA	2.2
1	A	300	GLU	2.2
1	A	754	ASP	2.2
1	A	632	LYS	2.2
1	A	745	LEU	2.2
1	A	481	TYR	2.1
1	A	296	ALA	2.1
1	A	145	LEU	2.1
1	A	305	LEU	2.1
1	A	149	GLY	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	399	ASN	2.1
1	A	567	ILE	2.1
1	A	469	ALA	2.1
1	A	285	GLN	2.1
1	A	252	VAL	2.1
1	A	586	GLY	2.1
1	A	642	ILE	2.0

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 monosaccharides 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.