Full wwPDB X-ray Structure Validation Report ⓘ

Dec 22, 2020 – 02:03 pm GMT

PDB ID	:	7B07
Title	:	TgoT_6G12 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	:	3.10 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:

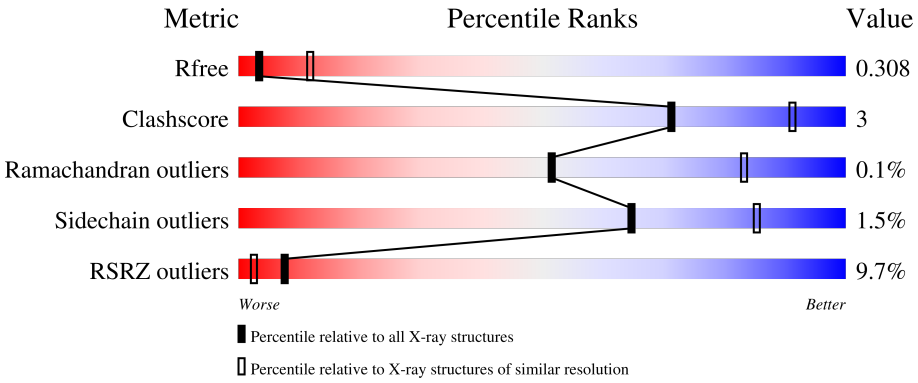
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.16
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.16

1 Overall quality at a glance (i)

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}	130704	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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	A	773	9% 88% 8% .

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6154 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
1	A	749	Total	C	N	O	S	0	0	0
			6150	3969	1036	1131	14			

There are 18 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	589	ALA	VAL	conflict	UNP P56689
A	609	LYS	GLU	conflict	UNP P56689
A	610	MET	ILE	conflict	UNP P56689
A	659	GLN	LYS	conflict	UNP P56689
A	664	GLN	GLU	conflict	UNP P56689
A	665	PRO	GLN	conflict	UNP P56689
A	668	LYS	ARG	conflict	UNP P56689
A	669	GLN	ASP	conflict	UNP P56689
A	671	HIS	LYS	conflict	UNP P56689
A	674	ARG	LYS	conflict	UNP P56689
A	676	ARG	THR	conflict	UNP P56689
A	681	SER	ALA	conflict	UNP P56689
A	704	PRO	LEU	conflict	UNP P56689
A	730	GLY	GLU	conflict	UNP P56689

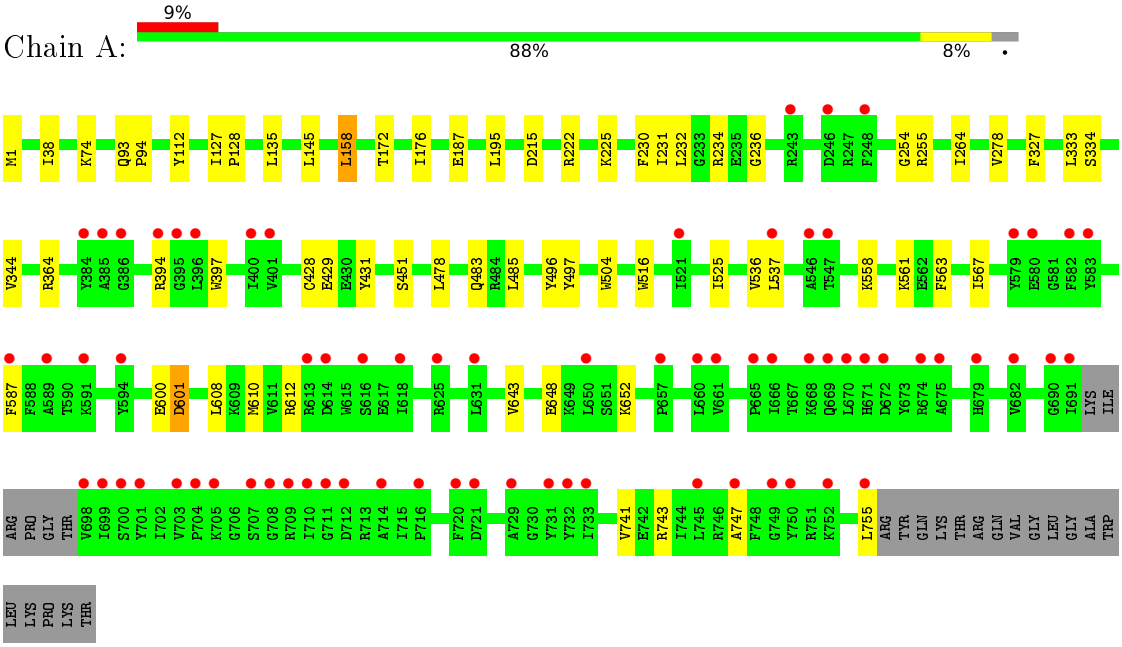
- Molecule 2 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	4	Total	Ca	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



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	71.89Å 110.49Å 111.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.53 – 3.10 49.53 – 3.10	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.53-3.10) 99.7 (49.53-3.10)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.51 (at 3.12Å)	Xtriage
Refinement program	BUSTER 2.10.3 (20-MAY-2020)	Depositor
R, R_{free}	0.262 , 0.265 0.296 , 0.308	Depositor DCC
R_{free} test set	836 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	100.3	Xtriage
Anisotropy	0.549	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 51.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.047 for -h,l,k	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	6154	wwPDB-VP
Average B, all atoms (Å^2)	140.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.12% 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: CA

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	0.22	0/6291	0.42	0/8484

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

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	6150	0	6180	43	0
2	A	4	0	0	0	0
All	All	6154	0	6180	43	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 (43) 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:394:ARG:HB3	1:A:747:ALA:HB1	1.77	0.66
1:A:558:LYS:HA	1:A:561:LYS:NZ	2.10	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:397:TRP:CE2	1:A:537:LEU:HD22	2.33	0.63
1:A:231:ILE:HG23	1:A:236:GLY:HA2	1.82	0.61
1:A:558:LYS:HA	1:A:561:LYS:HZ2	1.64	0.60
1:A:608:LEU:HG	1:A:743:ARG:HD2	1.84	0.60
1:A:187:GLU:OE2	1:A:222:ARG:HD2	2.01	0.60
1:A:333:LEU:HD13	1:A:485:LEU:HD22	1.87	0.56
1:A:145:LEU:HB2	1:A:158:LEU:HD11	1.89	0.55
1:A:135:LEU:HD12	1:A:327:PHE:HZ	1.73	0.53
1:A:364:ARG:NH2	1:A:451:SER:OG	2.41	0.53
1:A:608:LEU:HG	1:A:743:ARG:NE	2.23	0.53
1:A:264:ILE:HD13	1:A:278:VAL:HG11	1.90	0.53
1:A:587:PHE:CZ	1:A:747:ALA:HB3	2.44	0.52
1:A:525:ILE:HG23	1:A:536:VAL:HG21	1.91	0.52
1:A:38:ILE:HD11	1:A:93:GLN:HB3	1.92	0.52
1:A:397:TRP:CD1	1:A:537:LEU:HD22	2.45	0.51
1:A:648:GLU:HG2	1:A:652:LYS:HE2	1.92	0.51
1:A:608:LEU:HG	1:A:743:ARG:CD	2.40	0.51
1:A:587:PHE:HZ	1:A:747:ALA:HB3	1.76	0.50
1:A:38:ILE:HG12	1:A:112:TYR:HA	1.93	0.49
1:A:397:TRP:NE1	1:A:537:LEU:HD22	2.26	0.49
1:A:172:THR:HG21	1:A:176:ILE:HD12	1.95	0.48
1:A:231:ILE:HG22	1:A:231:ILE:O	2.13	0.47
1:A:232:LEU:O	1:A:255:ARG:NH1	2.47	0.47
1:A:563:PHE:O	1:A:567:ILE:HG13	2.14	0.47
1:A:231:ILE:CG2	1:A:236:GLY:HA2	2.47	0.45
1:A:234:ARG:HB2	1:A:254:GLY:HA3	1.98	0.45
1:A:643:VAL:HG21	1:A:741:VAL:HG11	1.98	0.45
1:A:334:SER:HA	1:A:344:VAL:HG21	1.99	0.45
1:A:428:CYS:HB2	1:A:431:TYR:CZ	2.51	0.45
1:A:600:GLU:O	1:A:601:ASP:HB2	2.18	0.43
1:A:1:MET:HE2	1:A:135:LEU:HD21	2.00	0.41
1:A:397:TRP:CE2	1:A:537:LEU:CD2	3.00	0.41
1:A:135:LEU:HD12	1:A:327:PHE:CZ	2.55	0.41
1:A:397:TRP:CD2	1:A:537:LEU:HD22	2.55	0.41
1:A:93:GLN:HB2	1:A:94:PRO:HD3	2.03	0.41
1:A:496:TYR:HE2	1:A:504:TRP:HB2	1.85	0.41
1:A:74:LYS:HB2	1:A:74:LYS:HE3	1.93	0.41
1:A:397:TRP:CD2	1:A:537:LEU:CD2	3.04	0.41
1:A:608:LEU:HD12	1:A:610:MET:HG3	2.03	0.41
1:A:195:LEU:HD21	1:A:230:PHE:CD1	2.57	0.40
1:A:127:ILE:HA	1:A:128:PRO:HD3	1.99	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	745/773 (96%)	723 (97%)	21 (3%)	1 (0%)	51 83

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	601	ASP

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	648/668 (97%)	638 (98%)	10 (2%)	65 85

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	158	LEU
1	A	215	ASP
1	A	225	LYS
1	A	429	GLU
1	A	478	LEU
1	A	483	GLN
1	A	497	TYR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	516	TRP
1	A	612	ARG
1	A	755	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	461	GLN

5.3.3 RNA

There are no RNA molecules in this entry.

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 monosaccharides in this entry.

5.6 Ligand geometry

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	A	749/773 (96%)	0.50	73 (9%) 7 2	91, 130, 207, 280	0

All (73) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	671	HIS	10.4
1	A	670	LEU	7.8
1	A	704	PRO	7.7
1	A	720	PHE	7.2
1	A	385	ALA	6.5
1	A	669	GLN	5.8
1	A	674	ARG	5.5
1	A	750	TYR	5.5
1	A	675	ALA	5.5
1	A	709	ARG	5.3
1	A	708	GLY	5.2
1	A	710	ILE	5.1
1	A	587	PHE	5.0
1	A	657	PRO	5.0
1	A	668	LYS	4.9
1	A	703	VAL	4.8
1	A	579	TYR	4.7
1	A	700	SER	4.5
1	A	699	ILE	4.3
1	A	698	VAL	4.1
1	A	701	TYR	3.9
1	A	395	GLY	3.8
1	A	547	THR	3.7
1	A	672	ASP	3.5
1	A	394	ARG	3.5
1	A	583	TYR	3.4
1	A	747	ALA	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	707	SER	3.4
1	A	386	GLY	3.3
1	A	711	GLY	3.3
1	A	614	ASP	3.3
1	A	521	ILE	3.3
1	A	243	ARG	3.2
1	A	613	ARG	3.1
1	A	384	TYR	3.1
1	A	731	TYR	3.1
1	A	401	VAL	3.1
1	A	714	ALA	3.1
1	A	537	LEU	3.1
1	A	546	ALA	3.0
1	A	591	LYS	2.9
1	A	400	ILE	2.8
1	A	631	LEU	2.8
1	A	660	LEU	2.7
1	A	246	ASP	2.7
1	A	729	ALA	2.7
1	A	712	ASP	2.7
1	A	716	PRO	2.7
1	A	618	ILE	2.6
1	A	589	ALA	2.6
1	A	594	TYR	2.6
1	A	665	PRO	2.6
1	A	650	LEU	2.5
1	A	732	TYR	2.5
1	A	721	ASP	2.5
1	A	749	GLY	2.4
1	A	582	PHE	2.4
1	A	745	LEU	2.4
1	A	396	LEU	2.4
1	A	625	ARG	2.3
1	A	679	HIS	2.2
1	A	690	GLY	2.2
1	A	682	VAL	2.2
1	A	755	LEU	2.2
1	A	691	ILE	2.2
1	A	752	LYS	2.1
1	A	666	ILE	2.1
1	A	580	GLU	2.1
1	A	705	LYS	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	616	SER	2.0
1	A	733	ILE	2.0
1	A	248	PHE	2.0
1	A	661	VAL	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)

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
2	CA	A	802	1/1	0.57	0.32	105,105,105,105	0
2	CA	A	803	1/1	0.83	0.23	124,124,124,124	0
2	CA	A	804	1/1	0.94	0.10	123,123,123,123	0
2	CA	A	801	1/1	0.97	0.44	85,85,85,85	0

6.5 Other polymers (i)

There are no such residues in this entry.