



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 6, 2026 – 01:37 PM UTC

PDB ID : 7OU3 / pdb_00007ou3
Title : Crystal structure of Tga-AGOG, an 8-oxoguanine DNA glycosylase from *Thermococcus gammatolerans*
Authors : Coste, F.; Confalonieri, F.; Castaing, B.
Deposited on : 2021-06-11
Resolution : 1.49 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

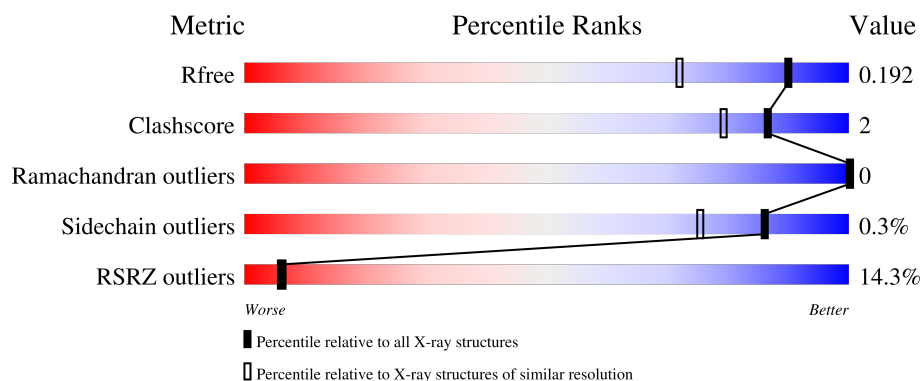
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.49 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}	180053	4037 (1.50-1.50)
Clashscore	190562	4235 (1.50-1.50)
Ramachandran outliers	187476	4153 (1.50-1.50)
Sidechain outliers	187428	4150 (1.50-1.50)
RSRZ outliers	180081	4039 (1.50-1.50)

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	266	11% 86% 7% 7%
1	B	266	16% 88% • 9%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4237 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 N-glycosylase/DNA lyase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	248	Total	C	N	O	S	0	5	0
			1986	1285	338	357	6			
1	B	241	Total	C	N	O	S	0	5	0
			1918	1240	322	351	5			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP C5A7E3
A	-1	SER	-	expression tag	UNP C5A7E3
A	0	HIS	-	expression tag	UNP C5A7E3
B	-2	GLY	-	expression tag	UNP C5A7E3
B	-1	SER	-	expression tag	UNP C5A7E3
B	0	HIS	-	expression tag	UNP C5A7E3

- Molecule 2 is CHLORIDE ION (CCD ID: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Cl	0	0
			1	1		
2	B	1	Total	Cl	0	0
			1	1		

- Molecule 3 is GLYCEROL (CCD ID: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		

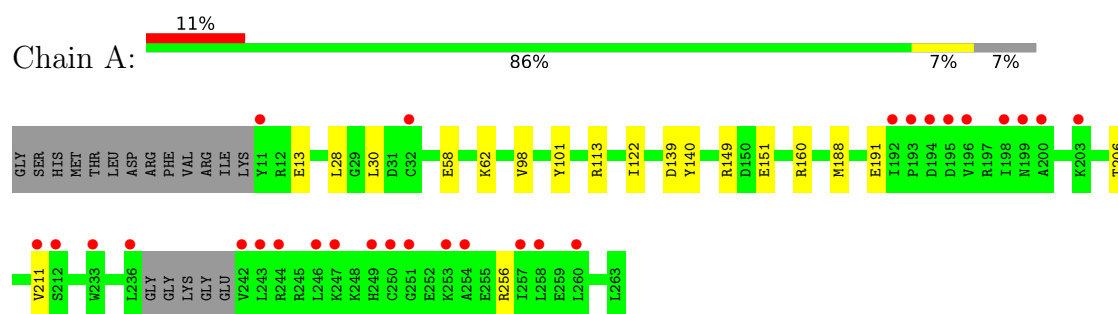
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	167	Total	O	0	0
			167	167		
4	B	152	Total	O	0	0
			152	152		

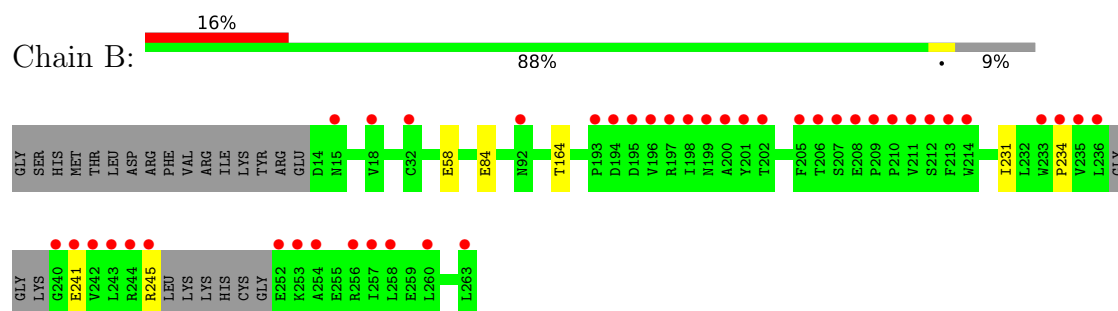
3 Residue-property plots

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: N-glycosylase/DNA lyase



• Molecule 1: N-glycosylase/DNA lyase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	61.43Å 62.45Å 139.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.15 – 1.49 46.15 – 1.49	Depositor EDS
% Data completeness (in resolution range)	99.9 (46.15-1.49) 95.5 (46.15-1.49)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.12 (at 1.49Å)	Xtriage
Refinement program	PHENIX 1.19_4092	Depositor
R, R_{free}	0.172 , 0.192 0.172 , 0.192	Depositor DCC
R_{free} test set	4390 reflections (4.94%)	wwPDB-VP
Wilson B-factor (Å ²)	18.5	Xtriage
Anisotropy	0.503	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 43.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.017 for k,h,-l	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	4237	wwPDB-VP
Average B, all atoms (Å ²)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.96% 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: CL, GOL

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	0/2046	0.43	0/2775
1	B	0.25	0/1975	0.44	0/2679
All	All	0.23	0/4021	0.44	0/5454

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	1986	0	1937	11	0
1	B	1918	0	1853	5	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
3	B	12	0	16	1	0
4	A	167	0	0	1	0
4	B	152	0	0	2	0
All	All	4237	0	3806	16	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 (16) 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:B:58:GLU:OE1	4:B:401:HOH:O	2.05	0.75
1:A:28:LEU:HD21	1:A:256:ARG:HB2	1.80	0.64
1:A:191:GLU:HG3	1:A:211:VAL:HG13	1.80	0.62
1:A:58[B]:GLU:HG2	1:A:62:LYS:HE3	1.89	0.54
1:A:98:VAL:HG22	1:A:122[A]:ILE:HG22	1.91	0.53
1:A:140:TYR:OH	1:A:151:GLU:OE1	2.13	0.50
1:B:84:GLU:HB2	3:B:301:GOL:H2	1.95	0.48
1:A:149:ARG:HD3	1:A:160:ARG:O	2.14	0.47
1:B:231:ILE:O	1:B:234:PRO:HD2	2.16	0.46
1:A:13:GLU:HG3	1:A:206:THR:HB	1.97	0.45
1:B:164[A]:THR:HG22	4:B:429:HOH:O	2.17	0.45
1:A:113:ARG:HB3	4:A:430:HOH:O	2.18	0.42
1:B:241:GLU:O	1:B:245:ARG:HG3	2.19	0.42
1:A:188:MET:HE3	1:A:188:MET:HB3	1.89	0.42
1:A:30:LEU:HD23	1:A:30:LEU:HA	1.95	0.41
1:A:101:TYR:CG	1:A:122[A]:ILE:HD13	2.56	0.40

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	249/266 (94%)	243 (98%)	6 (2%)	0	100	100
1	B	240/266 (90%)	236 (98%)	4 (2%)	0	100	100
All	All	489/532 (92%)	479 (98%)	10 (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	202/235 (86%)	201 (100%)	1 (0%)	81	66
1	B	194/235 (83%)	194 (100%)	0	100	100
All	All	396/470 (84%)	395 (100%)	1 (0%)	86	75

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	139	ASP

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	49	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 2 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	$\# Z > 2$
3	GOL	B	302	-	5,5,5	0.91	0	5,5,5	0.95	0
3	GOL	B	301	-	5,5,5	1.19	0	5,5,5	0.88	0

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
3	GOL	B	302	-	-	0/4/4/4	-
3	GOL	B	301	-	-	0/4/4/4	-

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	301	GOL	1	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	248/266 (93%)	0.51	28 (11%)	10 10	12, 26, 62, 73	5 (2%)
1	B	241/266 (90%)	0.59	42 (17%)	4 3	13, 25, 70, 84	5 (2%)
All	All	489/532 (91%)	0.55	70 (14%)	6 6	12, 26, 67, 84	10 (2%)

All (70) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	11	TYR	4.8
1	A	250	CYS	4.3
1	B	263	LEU	4.3
1	B	193	PRO	4.2
1	B	210	PRO	4.1
1	A	243	LEU	4.1
1	A	249	HIS	3.9
1	A	196	VAL	3.9
1	B	241	GLU	3.8
1	B	254	ALA	3.8
1	B	240	GLY	3.7
1	A	246	LEU	3.6
1	A	236	LEU	3.6
1	B	242	VAL	3.6
1	B	236	LEU	3.5
1	A	254	ALA	3.5
1	B	211	VAL	3.4
1	B	197	ARG	3.3
1	B	198	ILE	3.3
1	B	196	VAL	3.3
1	B	235	VAL	3.3
1	B	207	SER	3.3
1	B	257	ILE	3.3
1	B	209	PRO	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	195	ASP	3.2
1	B	243	LEU	3.1
1	B	206	THR	3.1
1	B	245	ARG	3.0
1	B	200	ALA	3.0
1	A	251	GLY	3.0
1	A	194	ASP	2.9
1	A	257	ILE	2.9
1	A	242	VAL	2.9
1	B	213	PHE	2.9
1	B	212	SER	2.8
1	A	233	TRP	2.8
1	B	258	LEU	2.8
1	B	253	LYS	2.8
1	B	202	THR	2.7
1	B	92	ASN	2.7
1	A	192	ILE	2.7
1	A	32	CYS	2.7
1	A	212	SER	2.6
1	B	256	ARG	2.6
1	B	208	GLU	2.6
1	B	214	TRP	2.6
1	B	15	ASN	2.6
1	B	244	ARG	2.6
1	B	201	TYR	2.5
1	B	199	ASN	2.5
1	B	205	PHE	2.5
1	A	244	ARG	2.5
1	B	252	GLU	2.5
1	A	247	LYS	2.4
1	A	258	LEU	2.4
1	A	253	LYS	2.4
1	A	198	ILE	2.4
1	B	260	LEU	2.4
1	B	32	CYS	2.3
1	A	211	VAL	2.3
1	B	234	PRO	2.3
1	A	260	LEU	2.2
1	A	193	PRO	2.2
1	A	203	LYS	2.2
1	A	200	ALA	2.2
1	A	199	ASN	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	194	ASP	2.1
1	B	18	VAL	2.1
1	A	195	ASP	2.1
1	B	233	TRP	2.1

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 oligosaccharides 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, 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	B-factors($\text{\AA}^2$)	Q<0.9
3	GOL	B	301	6/6	0.80	0.16	27,32,35,42	0
3	GOL	B	302	6/6	0.87	0.13	29,31,33,38	0
2	CL	A	301	1/1	0.99	0.11	20,20,20,20	0
2	CL	B	303	1/1	0.99	0.04	17,17,17,17	0

6.5 Other polymers [i](#)

There are no such residues in this entry.