



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 10, 2026 – 09:43 AM UTC

PDB ID : 2PQ0 / pdb_00002pq0
Title : Crystal structure of Hypothetical protein (gk_1056) from *Geobacillus Kaustophilus* HTA426
Authors : Kanaujia, S.P.; Jeyakanthan, J.; Kavyashree, M.; Sekar, K.; Agari, Y.; Ebihara, A.; Kuramitsu, S.; Shinkai, A.; Shiro, Y.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on : 2007-05-01
Resolution : 2.60 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

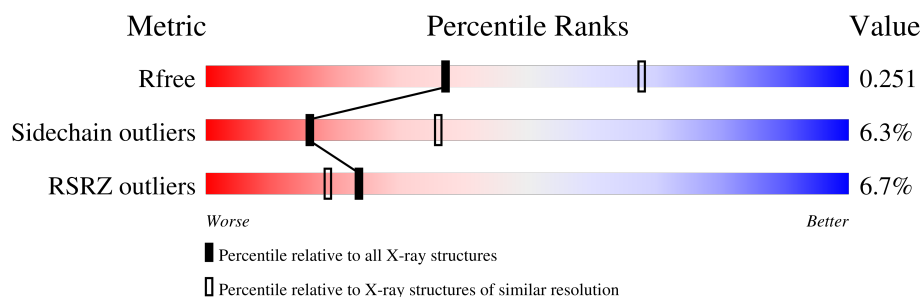
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.60 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	4008 (2.60-2.60)
Sidechain outliers	187428	4277 (2.60-2.60)
RSRZ outliers	180081	4008 (2.60-2.60)

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	258	7% 91% 8%
1	B	258	5% 94% 5%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4472 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 Hypothetical conserved protein GK1056.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	256	Total	C	N	O	S	Se	0	0	0
			2056	1321	355	371	1	8			
1	B	257	Total	C	N	O	S	Se	0	0	0
			2064	1326	356	372	1	9			

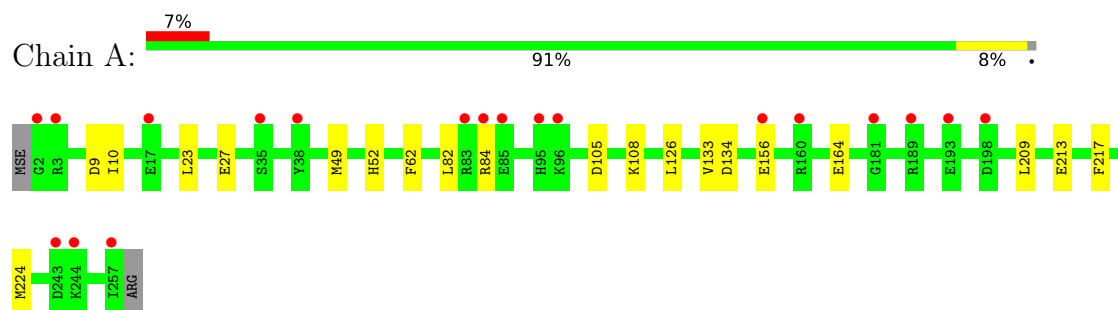
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	177	Total	O	0	0
			177	177		
2	B	175	Total	O	0	0
			175	175		

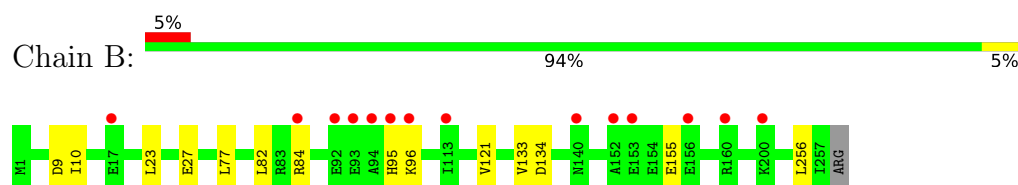
3 Residue-property plots [i](#)

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: Hypothetical conserved protein GK1056



- Molecule 1: Hypothetical conserved protein GK1056



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	114.05Å 114.05Å 182.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.07 – 2.60 39.07 – 2.60	Depositor EDS
% Data completeness (in resolution range)	95.4 (39.07-2.60) 97.5 (39.07-2.60)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.44 (at 2.61Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.208 , 0.241 0.219 , 0.251	Depositor DCC
R_{free} test set	1260 reflections (1.87%)	wwPDB-VP
Wilson B-factor (Å ²)	40.4	Xtriage
Anisotropy	0.326	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 53.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	4472	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.59% 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 [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	0.41	0/2097	0.93	11/2816 (0.4%)
1	B	0.41	0/2105	0.95	10/2826 (0.4%)
All	All	0.41	0/4202	0.94	21/5642 (0.4%)

There are no bond length outliers.

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	10	ILE	N-CA-C	12.83	122.42	110.74
1	A	133	VAL	N-CA-C	7.47	119.66	108.23
1	B	95	HIS	N-CA-C	-7.46	102.83	112.23
1	A	133	VAL	CB-CA-C	-7.37	102.05	110.96
1	A	9	ASP	N-CA-C	-7.05	100.03	110.46
1	B	133	VAL	N-CA-C	6.73	118.52	108.23
1	A	224	MSE	N-CA-C	6.54	120.19	109.72
1	B	9	ASP	N-CA-C	-6.47	100.94	110.52
1	A	10	ILE	N-CA-C	6.20	122.23	109.34
1	B	155	GLU	N-CA-C	6.06	118.85	111.82
1	B	256	LEU	N-CA-C	-5.98	105.47	112.89
1	B	134	ASP	CA-C-N	5.96	125.58	119.56
1	B	134	ASP	C-N-CA	5.96	125.58	119.56
1	A	217	PHE	N-CA-C	5.57	117.43	111.36
1	B	133	VAL	CB-CA-C	-5.24	104.62	110.96
1	A	108	LYS	N-CA-C	5.16	115.34	108.38
1	A	105	ASP	N-CA-C	-5.15	101.61	109.65
1	A	62	PHE	N-CA-C	5.09	116.26	108.42
1	B	121	VAL	N-CA-C	5.08	115.31	110.53
1	A	134	ASP	CA-C-N	5.06	124.85	119.28
1	A	134	ASP	C-N-CA	5.06	124.85	119.28

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

5.3 Torsion angles [i](#)

5.3.1 Protein backbone [i](#)

There are no protein backbone outliers to report in this entry.

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	208/211 (99%)	197 (95%)	11 (5%)	20	43
1	B	61/211 (29%)	55 (90%)	6 (10%)	7	17
All	All	269/422 (64%)	252 (94%)	17 (6%)	16	36

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	23	LEU
1	A	27	GLU
1	A	49	MSE
1	A	52	HIS
1	A	82	LEU
1	A	84	ARG
1	A	126	LEU
1	A	156	GLU
1	A	164	GLU
1	A	209	LEU
1	A	213	GLU
1	B	23	LEU
1	B	27	GLU
1	B	77	LEU
1	B	82	LEU
1	B	84	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	96	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	20	GLN
1	A	52	HIS
1	A	56	GLN
1	A	99	HIS
1	A	116	HIS
1	A	120	HIS
1	A	130	HIS
1	A	210	ASN
1	A	228	HIS
1	A	253	GLN
1	B	18	GLN
1	B	20	GLN
1	B	68	GLN
1	B	80	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 oligosaccharides 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	248/258 (96%)	0.32	19 (7%) 19 15	23, 41, 74, 90	0
1	B	248/258 (96%)	0.16	14 (5%) 30 24	21, 39, 72, 89	0
All	All	496/516 (96%)	0.24	33 (6%) 24 19	21, 40, 74, 90	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	84	ARG	5.2
1	B	160	ARG	4.9
1	A	2	GLY	4.7
1	A	156	GLU	3.9
1	B	113	ILE	3.8
1	B	95	HIS	3.8
1	A	193	GLU	3.8
1	B	94	ALA	3.4
1	A	17	GLU	3.4
1	B	140	ASN	3.4
1	A	160	ARG	3.3
1	A	83	ARG	3.3
1	A	257	ILE	3.1
1	A	189	ARG	3.0
1	B	200	LYS	2.9
1	B	152	ALA	2.9
1	A	243	ASP	2.8
1	A	38	TYR	2.8
1	B	84	ARG	2.6
1	A	96	LYS	2.6
1	B	156	GLU	2.6
1	B	92	GLU	2.4
1	B	96	LYS	2.3
1	B	153	GLU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	244	LYS	2.2
1	A	95	HIS	2.2
1	A	85	GLU	2.2
1	B	93	GLU	2.2
1	A	198	ASP	2.1
1	A	181	GLY	2.1
1	A	3	ARG	2.1
1	A	35	SER	2.0
1	B	17	GLU	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 oligosaccharides 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.