



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 08:12 pm GMT

PDB ID : 2PGS
Title : Crystal structure of a putative deoxyguanosinetriphosphate triphosphohydro-
lase from *Pseudomonas syringae* pv. *phaseolicola* 1448A
Authors : Rao, K.N.; Burley, S.K.; Swaminathan, S.; New York SGX Research Center
for Structural Genomics (NYSGXRC)
Deposited on : 2007-04-10
Resolution : 2.35 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 ⓘ 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	:	trunk30967
Percentile statistics	:	20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac	:	5.8.0158
CCP4	:	7.0 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk30967

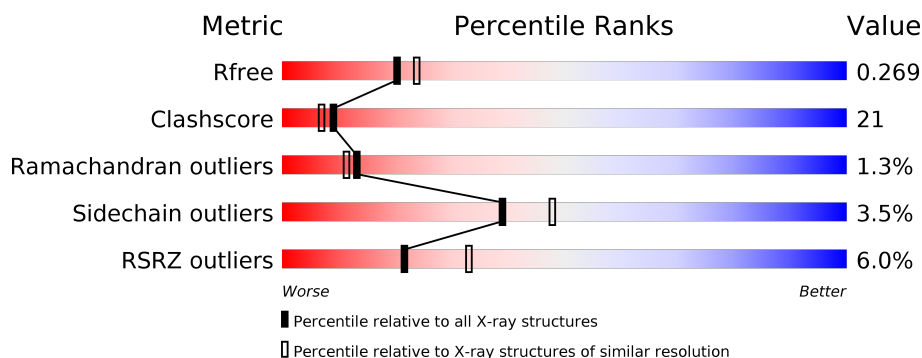
1 Overall quality at a glance

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}	111664	1015 (2.36-2.36)
Clashscore	122126	1081 (2.36-2.36)
Ramachandran outliers	120053	1066 (2.36-2.36)
Sidechain outliers	120020	1067 (2.36-2.36)
RSRZ outliers	108989	1002 (2.36-2.36)

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. 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	451	<div> <div>5%</div> <div>54%</div> <div>32%</div> <div>•</div> <div>13%</div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3105 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 Putative deoxyguanosinetriphosphate triphosphohydrolase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	393	Total	C	N	O	S	Se	0	0	0
			3056	1934	546	558	7	11			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	-	CLONING ARTIFACT	UNP Q48JX0
A	2	SER	-	CLONING ARTIFACT	UNP Q48JX0
A	79	MSE	MET	MODIFIED RESIDUE	UNP Q48JX0
A	100	MSE	MET	MODIFIED RESIDUE	UNP Q48JX0
A	140	MSE	MET	MODIFIED RESIDUE	UNP Q48JX0
A	239	MSE	MET	MODIFIED RESIDUE	UNP Q48JX0
A	258	MSE	MET	MODIFIED RESIDUE	UNP Q48JX0
A	334	MSE	MET	MODIFIED RESIDUE	UNP Q48JX0
A	348	MSE	MET	MODIFIED RESIDUE	UNP Q48JX0
A	419	MSE	MET	MODIFIED RESIDUE	UNP Q48JX0
A	426	MSE	MET	MODIFIED RESIDUE	UNP Q48JX0
A	434	MSE	MET	MODIFIED RESIDUE	UNP Q48JX0
A	438	MSE	MET	MODIFIED RESIDUE	UNP Q48JX0
A	444	GLU	-	CLONING ARTIFACT	UNP Q48JX0
A	445	GLY	-	CLONING ARTIFACT	UNP Q48JX0
A	446	HIS	-	CLONING ARTIFACT	UNP Q48JX0
A	447	HIS	-	CLONING ARTIFACT	UNP Q48JX0
A	448	HIS	-	CLONING ARTIFACT	UNP Q48JX0
A	449	HIS	-	CLONING ARTIFACT	UNP Q48JX0
A	450	HIS	-	CLONING ARTIFACT	UNP Q48JX0
A	451	HIS	-	CLONING ARTIFACT	UNP Q48JX0

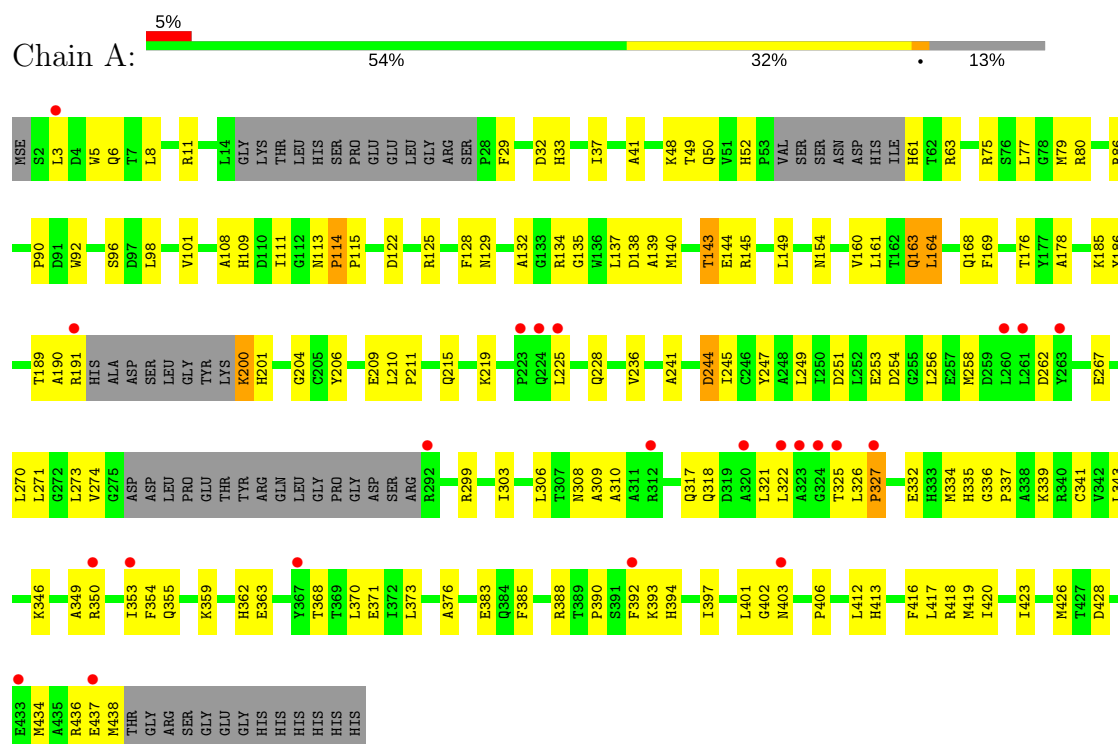
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	49	Total	O	0	0
			49	49		

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: Putative deoxyguanosinetriphosphate triphosphohydrolase



4 Data and refinement statistics

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, α , β , γ	103.73Å 103.73Å 159.12Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.31 – 2.35 49.31 – 2.33	Depositor EDS
% Data completeness (in resolution range)	97.7 (49.31-2.35) 97.1 (49.31-2.33)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.59 (at 2.32Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.251 , 0.273 0.251 , 0.269	Depositor DCC
R_{free} test set	619 reflections (2.78%)	wwPDB-VP
Wilson B-factor (Å ²)	39.2	Xtriage
Anisotropy	0.087	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 52.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	3105	wwPDB-VP
Average B, all atoms (Å ²)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.19% 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	1/3110 (0.0%)	0.60	0/4184

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	390	PRO	C-N	-10.79	1.09	1.34

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	3056	0	2982	129	0
2	A	49	0	0	2	0
All	All	3105	0	2982	129	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 21.

The worst 5 of 129 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:79:MSE:HE2	1:A:96:SER:HA	1.41	1.03

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:373:LEU:HD13	1:A:426:MSE:HE1	1.50	0.93
1:A:189:THR:HG22	1:A:191:ARG:H	1.35	0.90
1:A:317:GLN:O	1:A:321:LEU:HB2	1.79	0.83
1:A:273:LEU:HD23	1:A:341:CYS:SG	2.21	0.81

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	383/451 (85%)	359 (94%)	19 (5%)	5 (1%)	13	11

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	402	GLY
1	A	225	LEU
1	A	437	GLU
1	A	327	PRO
1	A	336	GLY

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	312/359 (87%)	301 (96%)	11 (4%)	39	48

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	200	LYS
1	A	228	GLN
1	A	355	GLN
1	A	164	LEU
1	A	262	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	154	ASN
1	A	362	HIS
1	A	308	ASN
1	A	146	ASN
1	A	163	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 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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	390:PRO	C	391:SER	N	1.09

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	382/451 (84%)	0.43	23 (6%)	22 33	17, 41, 65, 79	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	325	THR	4.7
1	A	392	PHE	4.4
1	A	312	ARG	3.5
1	A	191	ARG	3.1
1	A	323	ALA	2.9

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.