wwPDB X-ray Structure Validation Summary Report (i)

Sep 7, 2023 – 07:41 AM EDT

PDB ID	:	4GHK
Title	:	X-ray Crystal Structure of Gamma-glutamyl phosphate reductase from Burkholderia thailandensis
Authors	:	Seattle Structural Genomics Center for Infectious Disease (SSGCID)
Deposited on	:	2012-08-07
Resolution	:	2.25 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 (i) 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 (1)) were used in the production of this report:

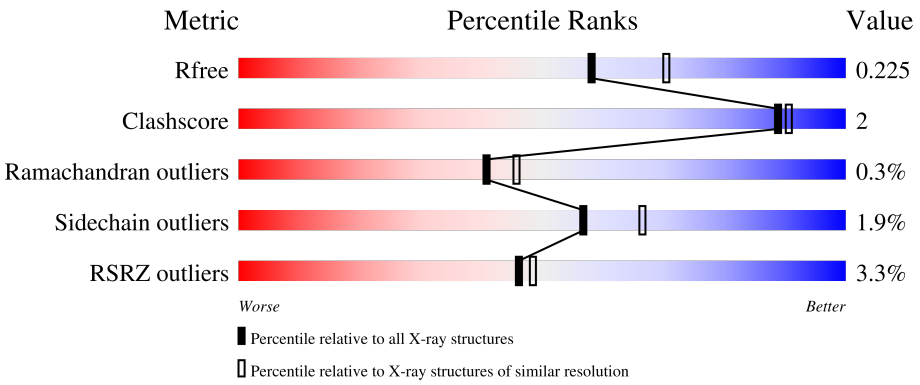
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.25 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	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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	444	4%, 86%, •, 9%
1	B	444	%, 87%, 5%, 8%
1	C	444	2%, 84%, 5%, 10%
1	D	444	5%, 84%, 6%, 10%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 12223 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 Gamma-glutamyl phosphate reductase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	404	2879	1787	519	559	14	0	1	0
1	B	410	2968	1842	537	576	13	0	1	0
1	C	398	2851	1775	517	546	13	0	0	0
1	D	400	2867	1784	523	546	14	0	1	0

There are 84 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-20	MET	-	expression tag	UNP Q2SZ88
A	-19	ALA	-	expression tag	UNP Q2SZ88
A	-18	HIS	-	expression tag	UNP Q2SZ88
A	-17	HIS	-	expression tag	UNP Q2SZ88
A	-16	HIS	-	expression tag	UNP Q2SZ88
A	-15	HIS	-	expression tag	UNP Q2SZ88
A	-14	HIS	-	expression tag	UNP Q2SZ88
A	-13	HIS	-	expression tag	UNP Q2SZ88
A	-12	MET	-	expression tag	UNP Q2SZ88
A	-11	GLY	-	expression tag	UNP Q2SZ88
A	-10	THR	-	expression tag	UNP Q2SZ88
A	-9	LEU	-	expression tag	UNP Q2SZ88
A	-8	GLU	-	expression tag	UNP Q2SZ88
A	-7	ALA	-	expression tag	UNP Q2SZ88
A	-6	GLN	-	expression tag	UNP Q2SZ88
A	-5	THR	-	expression tag	UNP Q2SZ88
A	-4	GLN	-	expression tag	UNP Q2SZ88
A	-3	GLY	-	expression tag	UNP Q2SZ88
A	-2	PRO	-	expression tag	UNP Q2SZ88
A	-1	GLY	-	expression tag	UNP Q2SZ88
A	0	SER	-	expression tag	UNP Q2SZ88

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-20	MET	-	expression tag	UNP Q2SZ88
B	-19	ALA	-	expression tag	UNP Q2SZ88
B	-18	HIS	-	expression tag	UNP Q2SZ88
B	-17	HIS	-	expression tag	UNP Q2SZ88
B	-16	HIS	-	expression tag	UNP Q2SZ88
B	-15	HIS	-	expression tag	UNP Q2SZ88
B	-14	HIS	-	expression tag	UNP Q2SZ88
B	-13	HIS	-	expression tag	UNP Q2SZ88
B	-12	MET	-	expression tag	UNP Q2SZ88
B	-11	GLY	-	expression tag	UNP Q2SZ88
B	-10	THR	-	expression tag	UNP Q2SZ88
B	-9	LEU	-	expression tag	UNP Q2SZ88
B	-8	GLU	-	expression tag	UNP Q2SZ88
B	-7	ALA	-	expression tag	UNP Q2SZ88
B	-6	GLN	-	expression tag	UNP Q2SZ88
B	-5	THR	-	expression tag	UNP Q2SZ88
B	-4	GLN	-	expression tag	UNP Q2SZ88
B	-3	GLY	-	expression tag	UNP Q2SZ88
B	-2	PRO	-	expression tag	UNP Q2SZ88
B	-1	GLY	-	expression tag	UNP Q2SZ88
B	0	SER	-	expression tag	UNP Q2SZ88
C	-20	MET	-	expression tag	UNP Q2SZ88
C	-19	ALA	-	expression tag	UNP Q2SZ88
C	-18	HIS	-	expression tag	UNP Q2SZ88
C	-17	HIS	-	expression tag	UNP Q2SZ88
C	-16	HIS	-	expression tag	UNP Q2SZ88
C	-15	HIS	-	expression tag	UNP Q2SZ88
C	-14	HIS	-	expression tag	UNP Q2SZ88
C	-13	HIS	-	expression tag	UNP Q2SZ88
C	-12	MET	-	expression tag	UNP Q2SZ88
C	-11	GLY	-	expression tag	UNP Q2SZ88
C	-10	THR	-	expression tag	UNP Q2SZ88
C	-9	LEU	-	expression tag	UNP Q2SZ88
C	-8	GLU	-	expression tag	UNP Q2SZ88
C	-7	ALA	-	expression tag	UNP Q2SZ88
C	-6	GLN	-	expression tag	UNP Q2SZ88
C	-5	THR	-	expression tag	UNP Q2SZ88
C	-4	GLN	-	expression tag	UNP Q2SZ88
C	-3	GLY	-	expression tag	UNP Q2SZ88
C	-2	PRO	-	expression tag	UNP Q2SZ88
C	-1	GLY	-	expression tag	UNP Q2SZ88
C	0	SER	-	expression tag	UNP Q2SZ88

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-20	MET	-	expression tag	UNP Q2SZ88
D	-19	ALA	-	expression tag	UNP Q2SZ88
D	-18	HIS	-	expression tag	UNP Q2SZ88
D	-17	HIS	-	expression tag	UNP Q2SZ88
D	-16	HIS	-	expression tag	UNP Q2SZ88
D	-15	HIS	-	expression tag	UNP Q2SZ88
D	-14	HIS	-	expression tag	UNP Q2SZ88
D	-13	HIS	-	expression tag	UNP Q2SZ88
D	-12	MET	-	expression tag	UNP Q2SZ88
D	-11	GLY	-	expression tag	UNP Q2SZ88
D	-10	THR	-	expression tag	UNP Q2SZ88
D	-9	LEU	-	expression tag	UNP Q2SZ88
D	-8	GLU	-	expression tag	UNP Q2SZ88
D	-7	ALA	-	expression tag	UNP Q2SZ88
D	-6	GLN	-	expression tag	UNP Q2SZ88
D	-5	THR	-	expression tag	UNP Q2SZ88
D	-4	GLN	-	expression tag	UNP Q2SZ88
D	-3	GLY	-	expression tag	UNP Q2SZ88
D	-2	PRO	-	expression tag	UNP Q2SZ88
D	-1	GLY	-	expression tag	UNP Q2SZ88
D	0	SER	-	expression tag	UNP Q2SZ88

- Molecule 2 is water.

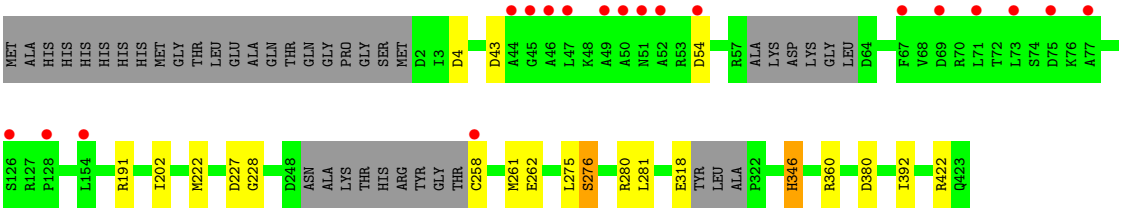
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	169	169	169	0	0
			Total	O		
2	B	182	182	182	0	0
			Total	O		
2	C	167	167	167	0	0
			Total	O		
2	D	140	140	140	0	0

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.

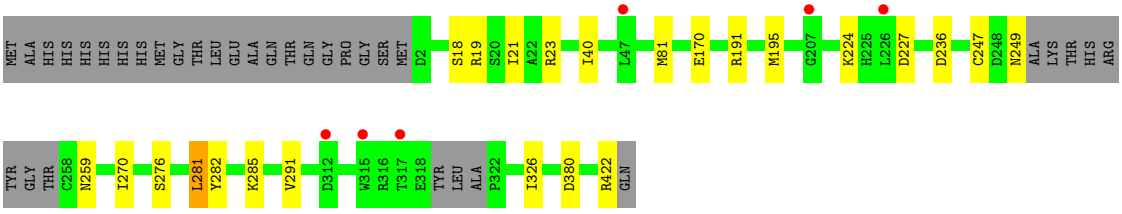
• Molecule 1: Gamma-glutamyl phosphate reductase

Chain A:



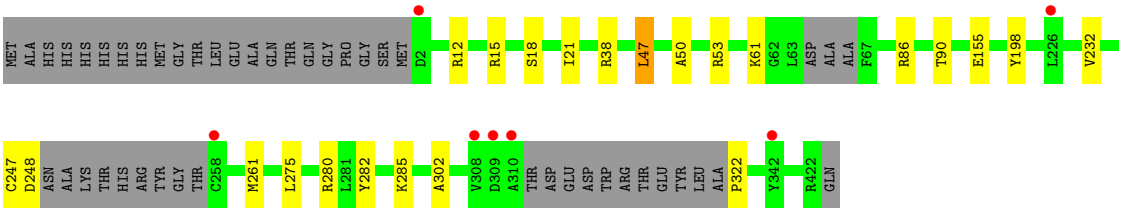
• Molecule 1: Gamma-glutamyl phosphate reductase

Chain B:



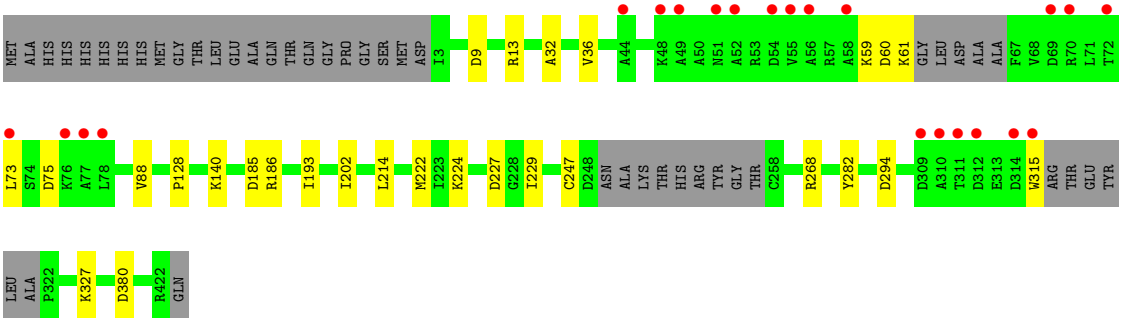
• Molecule 1: Gamma-glutamyl phosphate reductase

Chain C:



• Molecule 1: Gamma-glutamyl phosphate reductase

Chain D:


MET
ALA
HIS
HIS
HIS
HIS
HIS
MET
GLY
THR
LEU
GLU
ALA
GLN
THR
GLN
GLY
PRO
GLY
SER
MET
ASP
I3
D9
R13
A32
V36
A44
K48
A49
A50
N51
A52
R53
D54
V55
A56
R57
A58
K59
D60
K61
GLY
LEU
ASP
ALA
ALA
F67
V68
D69
R70
L71
T72
L73
S74
D75
K76
A77
L78
V88
P128
K140
D185
R186
I193
I202
L214
M222
I223
K224
D227
G228
I229
C247
D248
ASN
ALA
LYS
THR
HIS
ARG
TYR
GLY
THR
C258
R268
Y282
D294
D309
A310
T311
D312
E313
D314
W315
ARG
THR
GLU
TYR
LEU
ALA
P322
K327
D380
R422
GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	85.41Å 142.46Å 154.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	45.40 – 2.25 45.40 – 2.25	Depositor EDS
% Data completeness (in resolution range)	99.8 (45.40-2.25) 99.8 (45.40-2.25)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.50 (at 2.24Å)	Xtriage
Refinement program	PHENIX 1.8.1_1168	Depositor
R, R_{free}	0.180 , 0.225 0.182 , 0.225	Depositor DCC
R_{free} test set	4515 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å^2)	29.3	Xtriage
Anisotropy	0.535	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 46.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12223	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.41% 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 (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	0.35	0/2914	0.53	0/3962
1	B	0.35	0/3003	0.52	0/4078
1	C	0.35	0/2880	0.53	0/3906
1	D	0.34	0/2902	0.53	0/3941
All	All	0.35	0/11699	0.53	0/15887

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	2879	0	2785	8	0
1	B	2968	0	2940	11	0
1	C	2851	0	2833	13	0
1	D	2867	0	2821	12	0
2	A	169	0	0	0	0
2	B	182	0	0	3	0
2	C	167	0	0	1	0
2	D	140	0	0	0	0
All	All	12223	0	11379	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 2.

The worst 5 of 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:C:248:ASP:OD1	1:C:285:LYS:NZ	2.18	0.76
1:D:73:LEU:HD11	1:D:128:PRO:HG2	1.70	0.73
1:D:247:CYS:O	1:D:282:TYR:OH	2.11	0.68
1:A:276:SER:OG	1:A:280:ARG:NH2	2.31	0.63
1:C:38:ARG:NH1	2:C:629:HOH:O	2.30	0.62

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	397/444 (89%)	385 (97%)	12 (3%)	0	100	100
1	B	405/444 (91%)	396 (98%)	8 (2%)	1 (0%)	47	55
1	C	390/444 (88%)	380 (97%)	9 (2%)	1 (0%)	41	46
1	D	393/444 (88%)	381 (97%)	10 (2%)	2 (0%)	29	29
All	All	1585/1776 (89%)	1542 (97%)	39 (2%)	4 (0%)	41	46

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	61	LYS
1	D	60	ASP
1	B	227	ASP
1	D	227	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	275/344 (80%)	266 (97%)	9 (3%)	38	46
1	B	294/344 (86%)	288 (98%)	6 (2%)	55	64
1	C	278/344 (81%)	276 (99%)	2 (1%)	84	90
1	D	277/344 (80%)	273 (99%)	4 (1%)	67	76
All	All	1124/1376 (82%)	1103 (98%)	21 (2%)	57	66

5 of 21 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	380	ASP
1	D	75	ASP
1	D	380	ASP
1	D	224	LYS
1	C	322	PRO

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 (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	404/444 (90%)	-0.13	19 (4%)	31	34	18, 36, 76, 97	0
1	B	410/444 (92%)	-0.20	6 (1%)	73	75	19, 35, 70, 92	0
1	C	398/444 (89%)	-0.21	7 (1%)	68	71	20, 36, 68, 89	0
1	D	400/444 (90%)	-0.01	22 (5%)	25	27	20, 38, 86, 108	0
All	All	1612/1776 (90%)	-0.14	54 (3%)	46	48	18, 37, 77, 108	0

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	47	LEU	5.3
1	D	315	TRP	4.9
1	D	58	ALA	4.2
1	C	310	ALA	4.1
1	B	226	LEU	4.1

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 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.