



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 11:45 PM GMT

PDB ID : 1YJ8  
Title : Initial structural analysis of Plasmodium falciparum Glycerol-3-phosphate dehydrogenase  
Authors : Robien, M.A.; Hol, W.G.J.; Structural Genomics of Pathogenic Protozoa Consortium (SGPP)  
Deposited on : 2005-01-13  
Resolution : 2.85 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/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  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

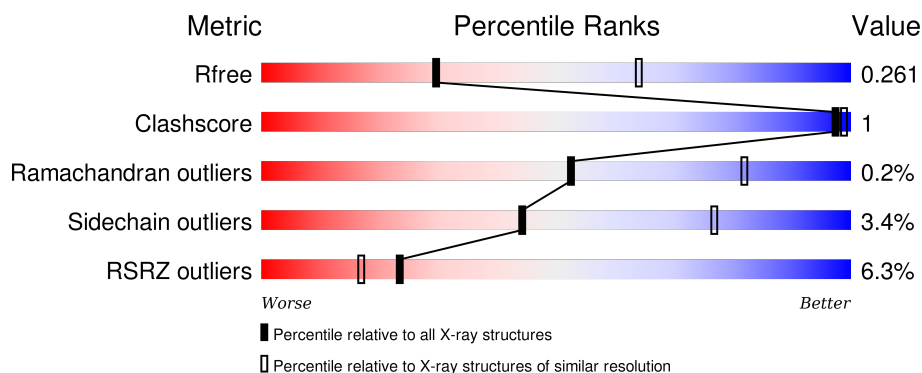
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.85 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))
$R_{free}$	91344	2228 (2.90-2.82)
Clashscore	102246	2499 (2.90-2.82)
Ramachandran outliers	100387	2439 (2.90-2.82)
Sidechain outliers	100360	2442 (2.90-2.82)
RSRZ outliers	91569	2236 (2.90-2.82)

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  $\geq 3$ , 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	375	<div> <div>2%</div> <div>92%</div> <div>5%</div> </div>
1	B	375	<div> <div>8%</div> <div>87%</div> <div>8%</div> <div>5%</div> </div>
1	C	375	<div> <div>8%</div> <div>89%</div> <div>6%</div> <div>5%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 8425 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 glycerol-3-phosphate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	357	Total	C	N	O	S	0	0	0
			2812	1806	466	522	18			
1	B	357	Total	C	N	O	S	0	0	0
			2806	1802	464	522	18			
1	C	357	Total	C	N	O	S	0	0	0
			2797	1796	462	521	18			

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	CLONING ARTIFACT	UNP Q8I5P5
A	2	ALA	-	CLONING ARTIFACT	UNP Q8I5P5
A	3	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
A	4	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
A	5	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
A	6	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
A	7	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
A	8	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
B	1	MET	-	CLONING ARTIFACT	UNP Q8I5P5
B	2	ALA	-	CLONING ARTIFACT	UNP Q8I5P5
B	3	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
B	4	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
B	5	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
B	6	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
B	7	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
B	8	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
C	1	MET	-	CLONING ARTIFACT	UNP Q8I5P5
C	2	ALA	-	CLONING ARTIFACT	UNP Q8I5P5
C	3	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
C	4	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
C	5	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
C	6	HIS	-	CLONING ARTIFACT	UNP Q8I5P5
C	7	HIS	-	CLONING ARTIFACT	UNP Q8I5P5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	8	HIS	-	CLONING ARTIFACT	UNP Q8I5P5

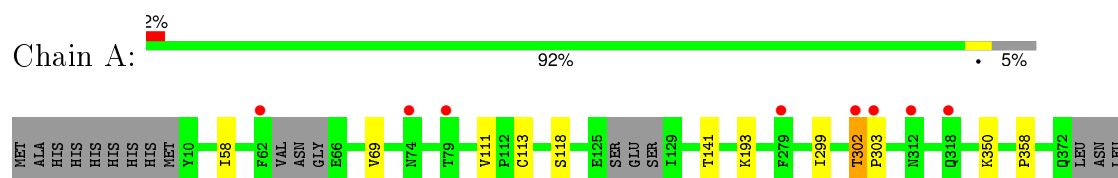
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	4	Total O 4 4	0	0
2	B	3	Total O 3 3	0	0
2	C	3	Total O 3 3	0	0

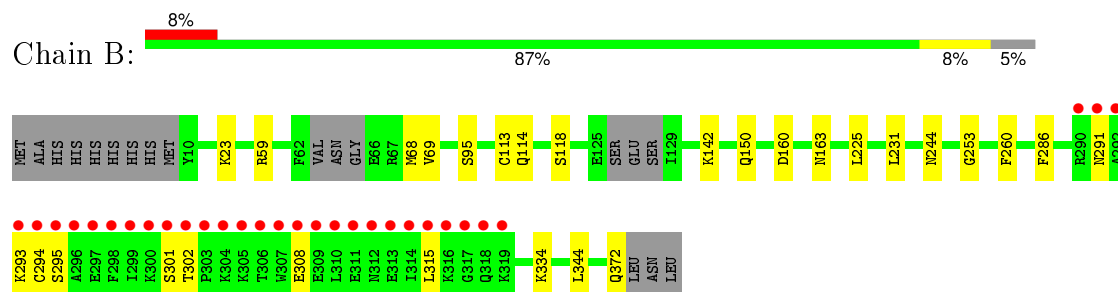
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

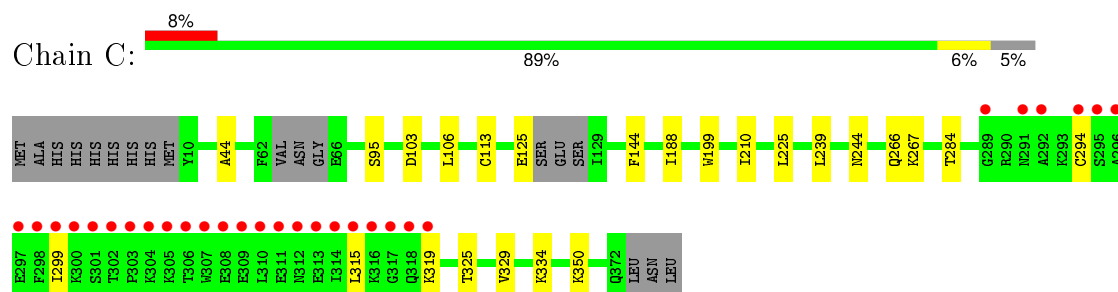
- Molecule 1: glycerol-3-phosphate dehydrogenase



- Molecule 1: glycerol-3-phosphate dehydrogenase



- Molecule 1: glycerol-3-phosphate dehydrogenase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 63 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	177.46 Å 177.46 Å 232.70 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	44.54 – 2.85 44.54 – 2.85	Depositor EDS
% Data completeness (in resolution range)	99.0 (44.54-2.85) 99.1 (44.54-2.85)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.13	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.93 (at 2.86 Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, $R_{free}$	0.236 , 0.256 0.240 , 0.261	Depositor DCC
$R_{free}$ test set	2557 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	58.7	Xtriage
Anisotropy	0.021	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 46.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 50515 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	8425	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	59.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.61% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.36	0/2862	0.50	0/3861
1	B	0.37	0/2856	0.50	0/3855
1	C	0.37	0/2847	0.49	0/3844
All	All	0.37	0/8565	0.50	0/11560

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	2812	0	2844	7	0
1	B	2806	0	2826	6	0
1	C	2797	0	2808	8	0
2	A	4	0	0	0	0
2	B	3	0	0	0	0
2	C	3	0	0	1	0
All	All	8425	0	8478	21	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 1.

All (21) 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:111:VAL:O	1:A:141:THR:HG21	2.02	0.59
1:A:302:THR:HB	1:A:303:PRO:HD3	1.85	0.59
1:A:58:ILE:HD11	1:A:69:VAL:HA	1.87	0.56
1:C:144:PHE:HE1	1:C:325:THR:HG22	1.73	0.53
1:C:294:CYS:SG	1:C:315:LEU:HG	2.49	0.52
1:A:302:THR:CB	1:A:303:PRO:HD3	2.41	0.51
1:B:69:VAL:HG21	1:B:95:SER:HB3	1.95	0.48
1:C:44:ALA:O	2:C:501:HOH:O	2.20	0.47
1:C:239:LEU:HG	1:C:299:ILE:HG21	1.97	0.45
1:A:302:THR:CB	1:A:303:PRO:CD	2.95	0.45
1:B:294:CYS:SG	1:B:315:LEU:HD21	2.57	0.45
1:B:253:GLY:HA2	1:B:344:LEU:HD23	1.99	0.44
1:C:188:ILE:HD12	1:C:210:ILE:HG21	2.00	0.43
1:B:225:LEU:HD11	1:B:260:PHE:CD2	2.53	0.43
1:A:111:VAL:C	1:A:141:THR:HG21	2.40	0.42
1:C:225:LEU:HD22	1:C:329:VAL:HG21	2.02	0.41
1:B:231:LEU:HD12	1:B:295:SER:HB3	2.02	0.41
1:B:68:MET:HE2	1:B:68:MET:HB2	1.98	0.41
1:C:144:PHE:CE1	1:C:325:THR:HG22	2.53	0.40
1:A:299:ILE:HD11	1:A:358:PRO:HB3	2.03	0.40
1:C:106:LEU:HB3	1:C:199:TRP:CH2	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	351/375 (94%)	342 (97%)	8 (2%)	1 (0%)	46	76
1	B	351/375 (94%)	338 (96%)	12 (3%)	1 (0%)	46	76
1	C	351/375 (94%)	341 (97%)	10 (3%)	0	100	100
All	All	1053/1125 (94%)	1021 (97%)	30 (3%)	2 (0%)	52	82



All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	302	THR
1	B	291	ASN

### 5.3.2 Protein sidechains ⓘ

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	314/333 (94%)	310 (99%)	4 (1%)	76	92
1	B	312/333 (94%)	295 (95%)	17 (5%)	27	59
1	C	310/333 (93%)	299 (96%)	11 (4%)	43	76
All	All	936/999 (94%)	904 (97%)	32 (3%)	44	77

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	113	CYS
1	A	118	SER
1	A	193	LYS
1	A	350	LYS
1	B	23	LYS
1	B	59	ARG
1	B	113	CYS
1	B	114	GLN
1	B	118	SER
1	B	142	LYS
1	B	150	GLN
1	B	160	ASP
1	B	163	ASN
1	B	244	ASN
1	B	286	PHE
1	B	293	LYS
1	B	301	SER
1	B	302	THR
1	B	308	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	334	LYS
1	B	372	GLN
1	C	95	SER
1	C	103	ASP
1	C	113	CYS
1	C	125	GLU
1	C	244	ASN
1	C	266	GLN
1	C	267	LYS
1	C	284	THR
1	C	319	LYS
1	C	334	LYS
1	C	350	LYS

Some 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 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 ⓘ

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	357/375 (95%)	0.12	8 (2%) 65 61	41, 55, 78, 92	2 (0%)
1	B	357/375 (95%)	0.29	30 (8%) 14 9	39, 56, 105, 113	0
1	C	357/375 (95%)	0.40	29 (8%) 15 9	39, 56, 108, 115	0
All	All	1071/1125 (95%)	0.27	67 (6%) 23 17	39, 55, 91, 115	2 (0%)

All (67) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	302	THR	10.0
1	C	307	TRP	9.6
1	B	303	PRO	9.1
1	C	303	PRO	8.7
1	C	310	LEU	8.3
1	B	302	THR	7.6
1	C	301	SER	7.0
1	B	299	ILE	6.7
1	C	318	GLN	6.3
1	B	314	ILE	5.9
1	C	306	THR	5.7
1	B	310	LEU	5.6
1	C	304	LYS	5.5
1	B	300	LYS	5.4
1	C	314	ILE	5.4
1	B	294	CYS	5.3
1	B	292	ALA	5.3
1	C	312	ASN	5.3
1	C	309	GLU	5.3
1	B	318	GLN	5.2
1	B	309	GLU	5.0
1	C	311	GLU	5.0
1	C	299	ILE	5.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	317	GLY	4.9
1	B	307	TRP	4.9
1	B	301	SER	4.8
1	C	305	LYS	4.8
1	B	304	LYS	4.7
1	B	305	LYS	4.4
1	C	300	LYS	4.3
1	B	306	THR	4.1
1	C	315	LEU	4.1
1	C	292	ALA	4.1
1	C	308	GLU	4.1
1	B	312	ASN	3.8
1	B	319	LYS	3.7
1	B	313	GLU	3.6
1	B	308	GLU	3.6
1	B	293	LYS	3.5
1	C	294	CYS	3.5
1	C	296	ALA	3.5
1	C	316	LYS	3.4
1	C	319	LYS	3.4
1	A	302	THR	3.4
1	C	317	GLY	3.4
1	B	297	GLU	3.4
1	C	298	PHE	3.2
1	B	316	LYS	3.2
1	B	290	ARG	3.1
1	B	315	LEU	3.0
1	B	296	ALA	2.8
1	C	291	ASN	2.7
1	B	291	ASN	2.7
1	C	313	GLU	2.6
1	C	289	GLY	2.6
1	C	297	GLU	2.5
1	A	279	PHE	2.4
1	B	295	SER	2.4
1	A	79	THR	2.3
1	A	318	GLN	2.3
1	A	74	ASN	2.2
1	A	303	PRO	2.2
1	C	295	SER	2.2
1	A	312	ASN	2.1
1	B	311	GLU	2.1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	298	PHE	2.1
1	A	62	PHE	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 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.