



# wwPDB X-ray Structure Validation Summary Report

Mar 1, 2014 – 02:34 AM GMT

PDB ID : 3DOC  
Title : Crystal Structure of TrkA glyceraldehyde-3-phosphatedehydrogenase from *Brucella melitensis*  
Authors : Seattle Structural Genomics Center for Infectious Disease (SSGCID)  
Deposited on : 2008-07-03  
Resolution : 2.40 Å(reported)

This is a wwPDB validation summary 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/ValidationPDFNotes.html>

---

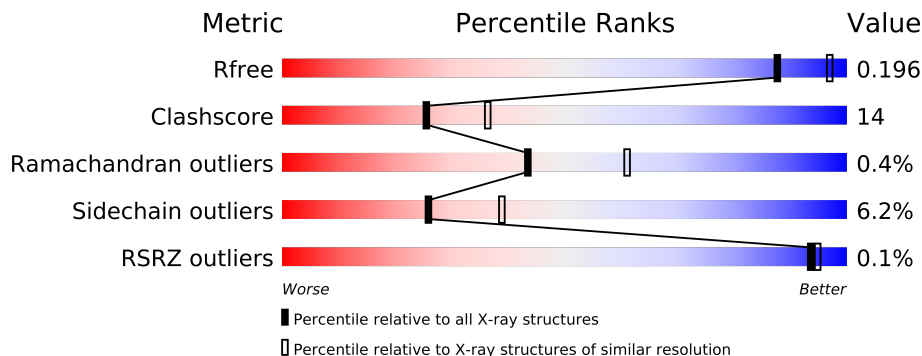
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.40 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))
$R_{free}$	66092	2207 (2.40-2.40)
Clashscore	79885	2789 (2.40-2.40)
Ramachandran outliers	78287	2736 (2.40-2.40)
Sidechain outliers	78261	2737 (2.40-2.40)
RSRZ outliers	66119	2210 (2.40-2.40)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	335	
1	B	335	
1	C	335	
1	D	335	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
2	NAD	A	901	-	X

## 2 Entry composition i

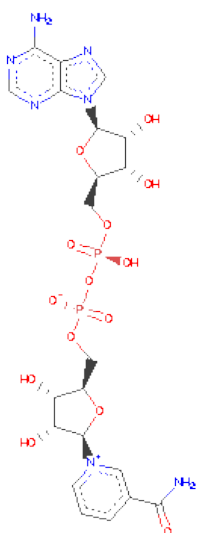
There are 3 unique types of molecules in this entry. The entry contains 10861 atoms, of which 0 are hydrogen and 0 are deuterium.

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 glyceraldehyde 3-phosphate dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	335	Total	C	N	O	S	0	0	0
			2543	1592	451	490	10			
1	B	335	Total	C	N	O	S	0	0	0
			2535	1587	450	488	10			
1	C	335	Total	C	N	O	S	0	0	0
			2547	1595	452	490	10			
1	D	334	Total	C	N	O	S	0	0	0
			2539	1590	451	489	9			

- Molecule 2 is NICOTINAMIDE-ADENINE-DINUCLEOTIDE (three-letter code: NAD) (formula:  $C_{21}H_{27}N_7O_{14}P_2$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	B	1	Total	C	N	O	P	0	0
			44	21	7	14	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	C	1	Total	C	N	O	P	0	0
			44	21	7	14	2		
2	D	1	Total	C	N	O	P	0	0
			44	21	7	14	2		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	121	Total	O	0	0
			121	121		
3	B	147	Total	O	0	0
			147	147		
3	C	139	Total	O	0	0
			139	139		
3	D	114	Total	O	0	0
			114	114		

### 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.

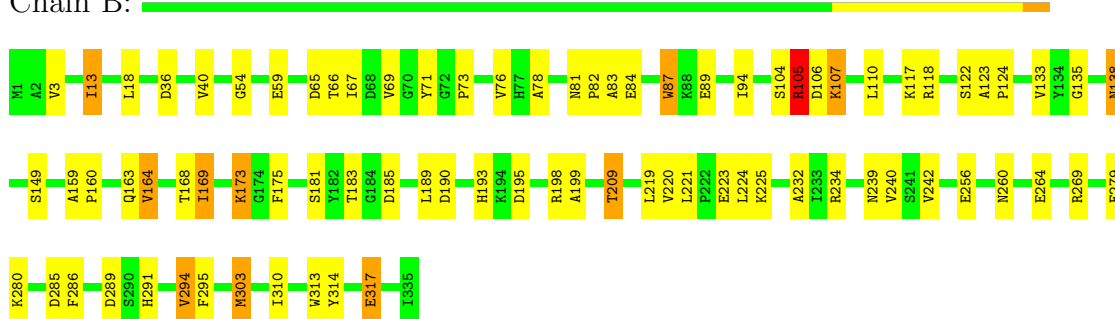
- Molecule 1: glyceraldehyde 3-phosphate dehydrogenase

Chain A:



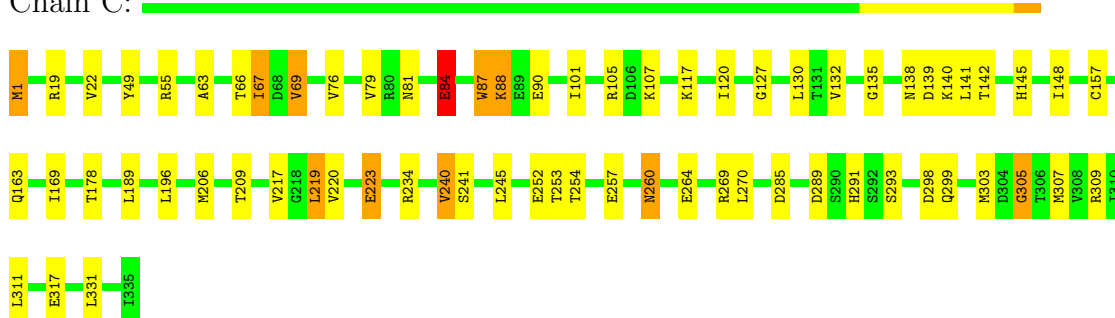
- Molecule 1: glyceraldehyde 3-phosphate dehydrogenase

Chain B:



- Molecule 1: glyceraldehyde 3-phosphate dehydrogenase

Chain C:



- Molecule 1: glyceraldehyde 3-phosphate dehydrogenase

Chain D:

	I248	A249	K250	R251	E252		E257		T277		D289		S290		H291		S292		S293		H296		T297		K301		V302		G305		V308		R309		W313		Y314		E317		D326		I335																																											
	Y134	G135		M138		T142		K143		D144		H145		S149		T155		M156		C157		L158		A159		P160		Q163		V164		T168		I169		E172		S181		Y182		T183		G184		D185		L189		D190		T191		L196		Y197		R198		T209		K215		A216		V217		V220		L221		A232		I233		R234		V240		S241		V242		V243		D244
	MET	A2		V5		L18		D28		D36		V40		Y49		G54		R55		F56		P57		A63		T66		I67		D68		Y71		G72		P73		I74		K75		V76		H77		A78		V79		R80		N81		P82		A83		E84		L85		P86		M87		E90		I94		R105		D106		A109		L110		H111		L112		K117		R118		G127

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	71.97Å 106.33Å 90.78Å 90.00° 107.89° 90.00°	Depositor
Resolution (Å)	50.00 – 2.40 45.28 – 2.41	Depositor EDS
% Data completeness (in resolution range)	96.9 (50.00-2.40) 97.0 (45.28-2.41)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.06 (at 2.39Å)	Xtriage
Refinement program	REFMAC 5.4.0067	Depositor
R, $R_{free}$	0.185 , 0.225 0.189 , 0.196	Depositor DCC
$R_{free}$ test set	2484 reflections (5.33%)	DCC
Wilson B-factor (Å <sup>2</sup> )	24.8	Xtriage
Anisotropy	0.076	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 17.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 49075 reflections	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	10861	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.94% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality i

### 5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: NAD

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	1.26	10/2586 (0.4%)	0.89	2/3515 (0.1%)
1	B	1.15	8/2578 (0.3%)	0.82	2/3506 (0.1%)
1	C	1.19	13/2590 (0.5%)	0.85	2/3519 (0.1%)
1	D	1.22	4/2582 (0.2%)	0.85	3/3509 (0.1%)
All	All	1.21	35/10336 (0.3%)	0.85	9/14049 (0.1%)

The worst 5 of 35 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	335	ILE	C-OXT	-14.41	0.95	1.23
1	A	90	GLU	CB-CG	-7.85	1.37	1.52
1	A	97	GLU	CD-OE1	-7.21	1.17	1.25
1	A	138	ASN	CB-CG	-6.73	1.35	1.51
1	A	197	TYR	CD1-CE1	-6.19	1.30	1.39

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	219	LEU	CA-CB-CG	9.27	136.62	115.30
1	B	105	ARG	NE-CZ-NH1	-6.69	116.95	120.30
1	D	68	ASP	CB-CG-OD1	6.63	124.27	118.30
1	D	244	ASP	CB-CG-OD1	6.37	124.03	118.30
1	C	19	ARG	NE-CZ-NH1	-6.20	117.20	120.30

There are no chirality outliers.

There are no planarity outliers.



## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2543	0	2551	102	0
1	B	2535	0	2534	73	0
1	C	2547	0	2562	47	0
1	D	2539	0	2550	64	0
2	A	44	0	26	20	0
2	B	44	0	26	2	0
2	C	44	0	26	0	0
2	D	44	0	26	0	0
3	A	121	0	0	3	0
3	B	147	0	0	8	0
3	C	139	0	0	2	0
3	D	114	0	0	1	0
All	All	10861	0	10301	281	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 14.

The worst 5 of 281 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:198:ARG:NH2	1:A:234:ARG:HH11	1.08	1.43
1:A:198:ARG:NH2	1:A:234:ARG:NH1	1.88	1.22
1:B:81:ASN:HD21	1:B:83:ALA:HB3	1.00	1.17
1:A:198:ARG:HH22	1:A:234:ARG:NH1	1.44	1.13
1:A:66:THR:HG21	1:A:73:PRO:HB3	1.30	1.11

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	333/335 (99%)	314 (94%)	18 (5%)	1 (0%)	50	68
1	B	333/335 (99%)	308 (92%)	24 (7%)	1 (0%)	50	68
1	C	333/335 (99%)	311 (93%)	20 (6%)	2 (1%)	33	47
1	D	332/335 (99%)	310 (93%)	21 (6%)	1 (0%)	50	68
All	All	1331/1340 (99%)	1243 (93%)	83 (6%)	5 (0%)	43	61

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	240	VAL
1	B	240	VAL
1	C	240	VAL
1	C	305	GLY
1	D	240	VAL

### 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	275/276 (100%)	259 (94%)	16 (6%)	28	43
1	B	273/276 (99%)	254 (93%)	19 (7%)	21	33
1	C	276/276 (100%)	261 (95%)	15 (5%)	31	47
1	D	275/276 (100%)	257 (94%)	18 (6%)	24	36
All	All	1099/1104 (100%)	1031 (94%)	68 (6%)	26	39

5 of 68 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	223	GLU
1	C	84	GLU
1	D	220	VAL
1	B	225	LYS
1	B	314	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 31 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	291	HIS
1	C	138	ASN
1	D	291	HIS
1	B	316	ASN
1	C	145	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

4 ligands are modelled in this entry.

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	$\# Z  > 2$
2	NAD	A	901	-	48,48,48	1.39	3 (6%)	73,73,73	1.69	7 (9%)
2	NAD	B	901	-	48,48,48	1.36	3 (6%)	73,73,73	1.67	11 (15%)
2	NAD	C	901	-	48,48,48	1.28	3 (6%)	73,73,73	1.86	10 (13%)
2	NAD	D	901	-	48,48,48	1.34	3 (6%)	73,73,73	1.84	10 (13%)

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
2	NAD	A	901	-	-	0/30/62/62	0/3/5/5
2	NAD	B	901	-	-	0/30/62/62	0/3/5/5
2	NAD	C	901	-	-	0/30/62/62	0/3/5/5
2	NAD	D	901	-	-	0/30/62/62	0/3/5/5

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	901	NAD	O7N-C7N	7.45	1.41	1.24
2	B	901	NAD	O7N-C7N	7.13	1.40	1.24
2	D	901	NAD	O7N-C7N	6.95	1.40	1.24
2	C	901	NAD	O7N-C7N	6.57	1.39	1.24
2	B	901	NAD	C2A-N3A	3.32	1.38	1.32

The worst 5 of 38 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	901	NAD	N3A-C2A-N1A	-11.08	119.44	128.71
2	D	901	NAD	N3A-C2A-N1A	-10.88	119.61	128.71
2	A	901	NAD	N3A-C2A-N1A	-10.42	120.00	128.71
2	B	901	NAD	N3A-C2A-N1A	-9.60	120.68	128.71
2	C	901	NAD	C3N-C7N-N7N	4.75	123.17	117.77

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers ⓘ

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	335/335 (100%)	-0.46	0 100 100	10, 21, 36, 44	0
1	B	335/335 (100%)	-0.44	0 100 100	9, 24, 37, 42	0
1	C	335/335 (100%)	-0.46	0 100 100	8, 21, 34, 41	0
1	D	334/335 (99%)	-0.45	0 100 100	10, 22, 34, 41	0
All	All	1339/1340 (99%)	-0.45	0 93 100	8, 22, 35, 44	0

There are no RSRZ outliers to report.

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

### 6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	NAD	A	901	44/44	0.28	3.85	68,72,75,75	0
2	NAD	B	901	44/44	0.10	-0.32	17,25,27,28	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	NAD	C	901	44/44	0.09	-0.40	12,17,19,20	0
2	NAD	D	901	44/44	0.09	-0.64	13,17,23,24	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.