

Full wwPDB X-ray Structure Validation Report (i)

Oct 31, 2023 – 10:13 AM JST

PDB ID : 5BMW

Title : Crystal structure of T75V mutant of Triosephosphate isomerase from Plas-

modium falciparum

Authors: Bandyopadhyay, D.; Murthy, M.R.N.; Balaram, H.; Balaram, P.

Deposited on : 2015-05-23

Resolution : 1.86 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 (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:

 $Mol Probity \quad : \quad 4.02b\text{--}467$

Mogul: 1.8.5 (274361), CSD as541be (2020)

Xtriage (Phenix) : 1.13 EDS : 2.36

Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)

 $Refmac \quad : \quad 5.8.0158$

CCP4 : 7.0.044 (Gargrove)

Ideal geometry (proteins) : Engh & Huber (2001) Ideal geometry (DNA, RNA) : Parkinson et al. (1996)

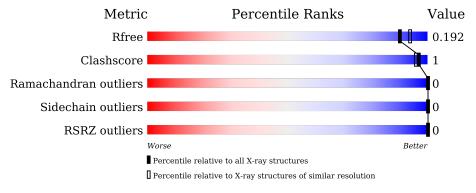
Validation Pipeline (wwPDB-VP) : 2.36

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 1.86 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	Similar resolution
Metric	$(\# \mathrm{Entries})$	$(\# ext{Entries}, ext{ resolution range}(ext{Å}))$
R_{free}	130704	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	248	97%				
1	В	248	94%	5%			



2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 4409 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.

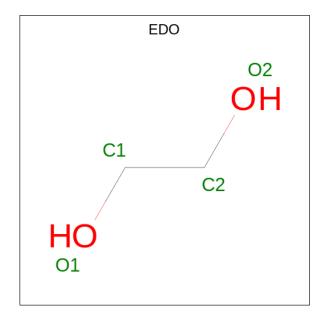
• Molecule 1 is a protein called Triosephosphate isomerase.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf	Trace		
1	Λ	246	Total	С	N	О	S	0	E	0
1	A	240	1953	1242	327	379	5	0	3	U
1	D	247	Total	С	N	О	S	0	1	0
1 B	247	1926	1225	323	373	5	0	1	0	

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	75	VAL	THR	engineered mutation	UNP Q07412
A	163	VAL	ALA	engineered mutation	UNP Q07412
В	75	VAL	THR	engineered mutation	UNP Q07412
В	163	VAL	ALA	engineered mutation	UNP Q07412

• Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



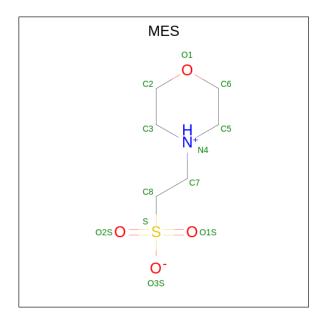


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	В	1	Total C O 4 2 2	0	0
2	В	1	Total C O 4 2 2	0	0

• Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Na 1 1	0	0
3	В	1	Total Na 1 1	0	0

• Molecule 4 is 2-(N-MORPHOLINO)-ETHANESULFONIC ACID (three-letter code: MES) (formula: $C_6H_{13}NO_4S$).





M	Iol	Chain	Residues	${f Atoms}$			ZeroOcc	AltConf		
	1	Λ	1	Total	С	N	О	S	0	0
- 1	4	A	1	12	6	1	4	1		

 \bullet Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Ca 1 1	0	0

• Molecule 6 is water.

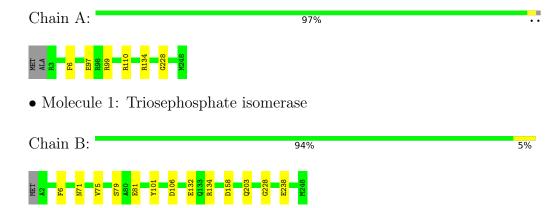
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	240	Total O 240 240	0	0
6	В	243	Total O 243 243	0	0



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: Triosephosphate isomerase





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants	39.21Å 76.02Å 74.36Å	Depositor
a, b, c, α , β , γ	90.00° 97.47° 90.00°	Depositor
Resolution (Å)	73.73 - 1.86	Depositor
rtesolution (A)	38.88 - 1.86	EDS
% Data completeness	99.7 (73.73-1.86)	Depositor
(in resolution range)	99.7 (38.88-1.86)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.06	Depositor
$< I/\sigma(I) > 1$	4.66 (at 1.85Å)	Xtriage
Refinement program	REFMAC 5.8.0103	Depositor
P. P.	0.131 , 0.182	Depositor
R, R_{free}	0.146 , 0.192	DCC
R_{free} test set	1816 reflections (4.99%)	wwPDB-VP
Wilson B-factor (Å ²)	11.7	Xtriage
Anisotropy	0.071	Xtriage
Bulk solvent $k_{sol}(e/Å^3)$, $B_{sol}(Å^2)$	0.37, 55.1	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	0.96	EDS
Total number of atoms	4409	wwPDB-VP
Average B, all atoms (Å ²)	12.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: The largest off-origin peak in the Patterson function is 7.31% of the height of the origin peak. No significant pseudotranslation is detected.

²Theoretical 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.



¹Intensities 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: NA, CA, EDO, MES

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	Bon RMSZ	nd lengths	Bond angles		
IVIOI	Mol Chain		# Z > 5	RMSZ	# Z >5	
1	A	0.97	0/2003	0.98	4/2704 (0.1%)	
1	В	0.96	1/1961 (0.1%)	0.92	4/2651 (0.2%)	
All	All	0.97	1/3964 (0.0%)	0.95	8/5355 (0.1%)	

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$\operatorname{Observed}(\text{\AA})$	$\operatorname{Ideal}(ext{\AA})$
1	В	81	GLU	CD-OE2	5.04	1.31	1.25

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$Observed(^o)$	$\mathbf{Ideal}(^o)$
1	A	110	ARG	NE-CZ-NH2	9.46	125.03	120.30
1	A	134	ARG	NE-CZ-NH2	9.13	124.86	120.30
1	A	134	ARG	NE-CZ-NH1	-8.51	116.05	120.30
1	A	99	ARG	NE-CZ-NH1	7.02	123.81	120.30
1	В	158	ASP	CB-CG-OD1	5.81	123.53	118.30
1	В	134	ARG	NE-CZ-NH1	-5.73	117.44	120.30
1	В	106	ASP	CB-CG-OD1	5.37	123.13	118.30
1	В	132	GLU	OE1-CD-OE2	-5.15	117.12	123.30

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 a	symmetric	unit.	whereas S	Svmm-	Clashes	lists s	vmmetr	v-related	clashes.
CIIC C	ob , militie of to	CLILIU,	11 11 C1 CCC K	O , 111111	CIGOTICE	TID OD D	, IIIIII OUI	, reracea	CICOLICO.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1953	0	1924	2	0
1	В	1926	0	1872	6	0
2	A	24	0	34	0	0
2	В	8	0	11	0	0
3	A	1	0	0	0	0
3	В	1	0	0	0	0
4	A	12	0	13	0	0
5	A	1	0	0	0	0
6	A	240	0	0	0	0
6	В	243	0	0	2	0
All	All	4409	0	3854	7	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 (7) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	$\begin{array}{c} {\rm Interatomic} \\ {\rm distance} \ ({\rm \AA}) \end{array}$	$\begin{array}{c} \text{Clash} \\ \text{overlap } (\text{\AA}) \end{array}$
1:B:101:TYR:CD1	1:B:101:TYR:CZ	2.48	0.93
1:B:203:GLN:CG	6:B:624:HOH:O	2.39	0.70
1:B:71:ASN:ND2	1:B:79:SER:OG	2.52	0.43
1:B:6:PHE:O	1:B:228:GLY:HA3	2.20	0.41
1:A:97:GLU:OE2	1:B:75:VAL:HG23	2.21	0.41
1:A:6:PHE:O	1:A:228:GLY:HA3	2.20	0.40
1:B:238:GLU:CG	6:B:522:HOH:O	2.69	0.40

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	Perce	ntiles
1	A	249/248 (100%)	245 (98%)	4 (2%)	0	100	100
1	В	246/248 (99%)	240 (98%)	6 (2%)	0	100	100
All	All	495/496 (100%)	485 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	Percent	iles
1	A	215/220 (98%)	215 (100%)	0	100 1	00
1	В	206/220 (94%)	206 (100%)	0	100 1	00
All	All	421/440 (96%)	421 (100%)	0	100 1	00

There are no protein residues with a non-rotameric sidechain to report.

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	50	HIS
1	В	71	ASN
1	В	159	ASN
1	В	200	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 3 are monoatomic - leaving 9 for Mogul analysis.

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	Вс	ond leng	ths	В	ond ang	les
MIOI	Type	Chain	nes	LILLK	Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	EDO	В	301	-	3,3,3	0.82	0	2,2,2	0.54	0
2	EDO	A	305	5	3,3,3	0.27	0	2,2,2	0.40	0
2	EDO	A	304	_	3,3,3	0.40	0	2,2,2	0.18	0
4	MES	A	308	-	12,12,12	1.15	1 (8%)	14,16,16	2.58	6 (42%)
2	EDO	A	301	-	3,3,3	0.10	0	2,2,2	0.44	0
2	EDO	A	303	-	3,3,3	1.10	0	2,2,2	0.31	0
2	EDO	В	302	-	3,3,3	0.21	0	2,2,2	0.92	0
2	EDO	A	302	-	3,3,3	0.54	0	2,2,2	0.60	0
2	EDO	A	306	-	3,3,3	0.41	0	2,2,2	0.37	0

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	EDO	В	301	-	-	1/1/1/1	-
2	EDO	A	305	5	-	0/1/1/1	-
2	EDO	A	304	-	-	0/1/1/1	-
4	MES	A	308	-	-	4/6/14/14	0/1/1/1
2	EDO	A	301	-	-	0/1/1/1	-
2	EDO	A	303	-	-	0/1/1/1	-
2	EDO	В	302	-	-	0/1/1/1	-
2	EDO	A	302	-	-	0/1/1/1	-
2	EDO	A	306	-	-	0/1/1/1	-



All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	\mathbf{Z}	$\operatorname{Observed}(\text{\AA})$	$\operatorname{Ideal}(ext{\AA})$
4	A	308	MES	C8-S	-3.45	1.72	1.77

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	$Observed(^o)$	$\operatorname{Ideal}({}^{o})$
4	A	308	MES	O1S-S-C8	6.93	115.25	106.92
4	A	308	MES	C6-C5-N4	3.20	114.95	110.10
4	A	308	MES	O3S-S-O1S	-3.13	103.63	111.27
4	A	308	MES	C6-O1-C2	2.71	118.94	109.89
4	A	308	MES	O2S-S-C8	-2.58	103.81	106.92
4	A	308	MES	C7-N4-C5	2.22	116.90	111.23

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	308	MES	C8-C7-N4-C5
4	A	308	MES	C7-C8-S-O2S
4	A	308	MES	C7-C8-S-O3S
4	A	308	MES	C7-C8-S-O1S
2	В	301	EDO	O1-C1-C2-O2

There are no ring outliers.

No monomer is involved in short contacts.

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, 95^{th} 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 $>$	$\#\mathrm{RSRZ}{>}2$		Z>2	$OWAB(A^2)$	Q<0.9
1	A	246/248 (99%)	-0.65	0 10	00	100	6, 10, 17, 29	0
1	В	247/248 (99%)	-0.68	0 10	00	100	6, 10, 18, 28	0
All	All	493/496 (99%)	-0.67	0 10	00	100	6, 10, 18, 29	0

There are no RSRZ outliers to report.

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)

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. The B-factors column lists the minimum, median, 95^{th} 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	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
2	EDO	В	302	4/4	0.89	0.17	14,15,15,19	4
2	EDO	A	301	4/4	0.93	0.18	11,12,12,13	4
2	EDO	A	303	4/4	0.95	0.13	7,7,7,7	4
2	EDO	В	301	4/4	0.95	0.16	4,5,5,5	4
2	EDO	A	302	4/4	0.95	0.07	17,21,21,24	0
2	EDO	A	306	4/4	0.96	0.10	15,17,18,19	0
4	MES	A	308	12/12	0.96	0.10	19,23,25,25	0

Continued on next page...



 $Continued\ from\ previous\ page...$

Mol	Type	Chain	Res	Atoms	RSCC	RSR	$\mathbf{B} ext{-}\mathbf{factors}(\mathbf{\mathring{A}}^2)$	Q < 0.9
2	EDO	A	304	4/4	0.97	0.06	10,11,12,14	0
2	EDO	A	305	4/4	0.98	0.07	9,9,9,9	0
3	NA	В	303	1/1	0.99	0.10	8,8,8,8	0
5	CA	A	309	1/1	0.99	0.06	10,10,10,10	0
3	NA	A	307	1/1	1.00	0.08	6,6,6,6	0

6.5 Other polymers (i)

There are no such residues in this entry.

